UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

Expression profiling of rust resistance-related genes in coffee and development of a virus induced gene silencing (VIGS)-based tool for functional analysis

Cíntia da Silva Macedo

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:

Doutora Dora Batista
Doutora Célia Miguel

2019

AGRADECIMENTOS

Na execução deste trabalho enfrentei inúmeros desafios e incertezas, mas também alegrias. Toda esta experiência me tornou numa pessoa mais resiliente e feliz com as minhas escolhas.
Concluir este capítulo da minha vida só foi possível com o apoio de várias pessoas, às quais dedico estes agradecimentos.

Em primeira instância, um agradecimento especial às minhas orientadoras, Doutora Dora Batista, a quem agradeço a excelente orientação, apoio e incentivo prestados mesmo nos períodos mais difíceis, assim como à Doutora Célia Miguel por sempre ter demonstrado imensa disponibilidade para me ajudar enquanto se aliava de uma enorme simpatia para comigo. A ambas agradeço o empenho e contribuição para o meu enriquecimento pessoal e académico.

Agradeço ainda às pessoas que me apoiaram e me guiaram na parte mais prática do trabalho, incluindo a espetacular equipa da Doutora Célia Miguel, onde todos se mostraram sempre disponíveis para me ajudar, e à fantástica equipa do CIFC, local onde fui extremamente bem recebida e que vou levar sempre no coração. Quero assim agradecer de forma mais particular à Doutora Maria do Céu, Doutora Andreia Loureiro, Engenheiro Vitor Várzea, Doutora Paula Pereira, Doutora Helena Azinheira, Doutora Leonor Guerra-Guimarães e não posso esquecer nem deixar de agradecer aos restantes funcionários do CIFC, pela disponibilidade, incentivo, sugestões e pelo excelente ambiente de trabalho que me proporcionaram. Em especial, agradeço à Inês Diniz por ter tido paciência, disponibilidade e pelo apoio técnico prestado no desenvolvimento deste trabalho.

Ao meu grupo de amigos, Joana, Bonnie, Luís e Tiago agradeço o apoio, compreensão, força e por me darem ouvidos quando me sinto mais em baixo. E agradeço igualmente às minhas duas colegas que já me acompanham desde a licenciatura, Mestre Nelly Silva e Mestre Mariana Oliveira, que sempre se mostraram disponíveis para me ajudar. A vocês estou grata pela nossa amizade.

Ao meu namorado, João Carvalho, pelo amor, companheirismo e apoio incondicional, agradeço também a compreensão, as brincadeiras e alegria que me proporcionou constantemente.

À minha avó, Isabel Macedo, e aos restantes familiares que sempre mostraram preocupação e ofereceram imenso apoio para que concretizasse esta etapa da minha vida.

E claro, aos meus queridos pais. À minha mãe, por me apoiar em todas as decisões e mostrar orgulho na pessoa que me tornei. E ao meu pai, que é o meu ídolo e a quem recorro nos piores e nos melhores momentos, pela força e otimismo que sempre me deu. Sem vocês nada disto seria possível.

Obrigada por tudo!

RESUMO

A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix*, é a doença mais importante do cafeeiro-Arábica, afetando praticamente todas as regiões do mundo onde se produz café. O controlo desta ferrugem pode ser conseguido recorrendo à luta química, nomeadamente através da aplicação de fungicidas cúpricos, como medida preventiva. No entanto, este meio de combate é desvantajoso tanto a nível económico como ambiental, além de que a sua aplicação nem sempre é rentável. Uma medida mais eficaz no combate a esta doença é a utilização de plantas de variedades resistentes à ferrugem. Embora o melhoramento genético do cafeeiro por métodos convencionais para o aumento desta resistência tenha tido muito sucesso, a sua durabilidade é constantemente desafiada pela alta adaptabilidade do fungo. Um melhor entendimento dos mecanismos moleculares de resistência do cafeeiro ao fungo é crucial para aumentar a eficácia das medidas de controlo da doença e prolongar o tempo de vida de cultivares comerciais resistentes. Estudos anteriores de proteómica e de expressão de genes identificaram diversas proteínas/genes putativamente envolvidos em vias de sinalização, reconhecimento e processos de defesa, associados à resposta das plantas de cafeeiro à infeção pelo fungo.

O objetivo deste trabalho foi investigar os perfis de expressão de genes putativamente envolvidos na resistência do cafeeiro a *H. vastatrix* e desenvolver uma ferramenta baseada na técnica virus induced gene silencing (VIGS) para posteriormente elucidar à cerca das funções de genes candidatos. Numa primeira parte do trabalho, plantas *Coffea arabica* (variedade Caturra) foram inoculadas com o isolado 71 (raça VI) de *H. vastatrix*, formando uma interação incompatível, e com o isolado 1427 (raça II) formando uma interação compatível. Posteriormente, foram colhidas folhas inoculadas 24, 48, 72 e 96 horas após inoculação (hpi). Para um estudo comparativo, foram também colhidas folhas controlo em cada tempo, as quais não sofreram qualquer tipo de tratamento. Com o intuito de monitorizar a evolução da doença e validar os tempos de infeção selecionados, foram feitas análises citológicas para observar os diferentes tipos de estruturas e respetiva abundância que se desenvolvem ao longo do processo de infeção, comparando entre interações compatíveis e incompatíveis. Em relação às fases de pré-penetração (24 hpi), não existe diferença significativa entre os dois tipos de interação, sendo que em ambas as respostas se observaram a germinação dos uredósporos e posterior diferenciação de apressórios. Na fase de pós- penetração, em ambos os tipos de interações existe o desenvolvimento das estruturas de infeção do fungo. No entanto, na interação incompatível, *H. vastatrix* cessou o seu crescimento mais frequentemente na fase de apressório, chegando a formar células mãe do haustório (CMH) com haustórios numa baixa percentagem de zonas de infeção (2% às 96 hpi), não progredindo para além deste ponto. Pelo contrário, na interação compatível, o crescimento do fungo prosseguiu sem aparente inibição com a formação de CMH com haustório a partir das 48 hpi, atingindo maior percentagem de zonas de infeção às 96 hpi e acabando por colonizar os tecidos foliares. Para as interações cafeeiro-ferrugem em estudo, foram analisados os perfis de expressão de sete genes putativamente envolvidos em mecanismos de reconhecimento, sinalização e defesa (*GL18058*, *GL22853*, *Asp23673*, *Asp203*, *PR1*, *PR10* e *RLK*), através de PCR quantitativo em tempo real. Os genes *GADPH* e *Ubiquitina* foram usados como genes de referência. Os genes *GL18058* e *Asp203* aparentemente não apresentam níveis de expressão significativos em nenhuma das interações nos tempos de infeção estudados, à exceção do gene *Asp203* na interação compatível que tem um pico de ativação às 72 hpi. De notar que no gene *GL18058* existe um perfil contrastante entre as duas interações às 24 hpi, e também uma alteração no perfil de expressão na interação incompatível, entre as 24 e as 48 hpi, sendo assim interessante estudar tempos intermédios deste intervalo em futuros trabalhos. Os genes *GL22853* e *Asp23673* mostraram estar reprimidos em todos os tempos de infeção estudados, com exceção das 96 hpi no gene *Asp23673* que apresenta uma baixa expressão. Para estes

dois genes, e apesar de não haver diferenças significativas na sua expressão comparando as duas interações, detetou-se diferença estatística entre alguns dos tempos estudados em ambas as interações. Por sua vez, o gene *PR1* tem um pico de ativação às 24 hpi na interação incompatível, sendo que nos restantes tempos está pouco ativo, sugerindo que a sua expressão poderá estar envolvida no bloqueio do desenvolvimento do fungo. Por outro lado, os genes *PR10* e *RLK* têm o seu maior pico de expressão às 48 hpi. Na interação compatível, nestes 3 genes, existe um pico de expressão às 72 hpi, correspondendo possivelmente a uma resposta à infeção, no entanto demasiado tardia para impedir o crescimento do fungo. Para nenhum dos genes analisados existe uma diferença estatisticamente relevante nos perfis de expressão, entre os dois tipos de interação. Hipoteticamente, os genes pouco ativos ou mesmo reprimidos podem ter tido um maior envolvimento no processo de infeção numa fase mais precoce, correspondendo a tempos que não foram estudados neste trabalho, ou podem ter sido alvo de uma regulação negativa para tentar inibir o desenvolvimento do fungo. Nos restantes genes que estão putativamente associados a processos de reconhecimento e defesa da planta, observa-se uma tendência para um aumento da expressão nos dois tipos de interação, sendo que na interação incompatível esse aumento dá-se numa fase anterior do processo de infeção. O mesmo tipo de resposta numa fase mais tardia da infeção poderá não ser eficaz para bloquear o crescimento do fungo resultando numa interação compatível, o que é consistente com as análises citológicas.

A transformação genética representa outra abordagem para caracterização funcional de genes candidatos que, apesar de muito eficiente nalgumas espécies é muitas vezes difícil de aplicar a grande parte das espécies de plantas lenhosas, incluindo o cafeeiro. Neste trabalho, procurou-se desenvolver uma metodologia eficaz, com base na tecnologia recombinante de vetores virais para o silenciamento de genes candidatos denominada VIGS, que leva à perda de função de proteínas alvo. Esta abordagem explora mecanismos de silenciamento pós transcricional (PTGS-post-transcriptional gene silencing) utilizados pelas plantas como defesa contra vírus invasores. Esta estratégia tem sido muito usada nos últimos anos, em várias plantas, para estudar a função de genes de interesse. O vetor viral é construído por inserção no genoma viral de um fragmento do gene a silenciar no hospedeiro que, neste caso, consiste num vetor derivado do Tobacco rattle virus (TRV), um vírus bipartido com dois genomas TRV1 e TRV2. Posteriormente, procede-se à transferência desta construção para a planta através de *Agrobacterium tumefaciens.* Assim é possível redirecionar o mecanismo de defesa mediado por RNA do hospedeiro para o silenciamento ou repressão da expressão do gene de interesse na planta, afetando a função da proteína correspondente. Neste trabalho, foi construído um vetor TRV2::PDS (Phytoene desaturase), utilizado como controlo positivo por ter um fragmento do gene repórter *PDS*, o qual foi transferido para cafeeiros e plantas de tabaco utilizadas como controlo positivo, através de *A. tumefaciens* por três métodos de inoculação distintos: agroinoculação, agrodrench e imersão das raízes das plantas na suspensão de bactérias contendo a construção. Foram testados diversos fatores que estão descritos como influenciadores da eficácia do sistema de silenciamento do gene *PDS*, tais como o estádio de desenvolvimento da planta, OD_{600} da suspensão bacteriana e temperatura pós inoculação. Após testadas as diversas condições, nenhum dos indivíduos, tanto de tabaco como cafeeiro, evidenciou cloroses nas folhas emergentes, que seria o fenótipo esperado após aplicação do sistema VIGS, em consequência do silenciamento do gene *PDS*. No entanto, foi possível detetar a presença do vetor viral no genoma de plantas inoculadas, mostrando que embora a integração do vetor pareça ter ocorrido, mesmo que de uma forma transiente ou pouco eficiente, por alguma razão não se verificou a repressão do gene *PDS*, ou alternativamente, a diminuição da expressão do gene não foi suficiente para mostrar o fenótipo esperado. Estes resultados confirmaram trabalhos anteriores e identificaram genes de interesse para analises funcionais futuras. Estudos adicionais são necessários para o desenvolvimento de um sistema VIGS otimizado em cafeeiro, o que seria uma ferramenta muito útil para a caracterização funcional de genes putativamente envolvidos na resistência da planta ao fungo.

Após validação da função dos genes candidatos e confirmação do envolvimento na resistência à ferrugem, este conhecimento seria importante para desenvolver marcadores moleculares que permitissem selecionar variedades resistentes de forma expedita.

Palavras-chave: ferrugem do cafeeiro, interação compatível, interação incompatível, agroinoculação, tabaco

ABSTRACT

Coffee leaf rust is a disease caused by the fungus *Hemileia vastatrix* that leads to big losses in coffee production. Management strategies used to combat the disease are not ideal or completely effective and therefore, increasing the knowledge on the molecular mechanisms involved in coffee-pathogen interactions is a key approach to support breeding programs for increased resistance. This work aimed to investigate the expression profiles of candidate genes for coffee resistance in compatible and incompatible interactions with *H. vastatrix*, and to develop a virus induced gene silencing (VIGS) system to further validate and elucidate the function of selected candidate genes. In the first part of this work, key-time points of the infection process of compatible and incompatible interactions between *Coffea arabica* (var. Caturra) and *H. vastatrix* were monitored by cytological observations to validate the time-course for subsequent gene expression analysis. Gene expression profiles of seven candidate genes (*GL18058*, *GL22853*, *Asp23673*, *Asp203*, *PR1*, *PR10* and *RLK*) putatively involved in signaling, recognition and defense pathways were analyzed by qPCR. Genes *GL18058* and *Asp203* seem to have no significant expression in the time-points selected. Genes *GL22853* and *Asp23673* are repressed throughout most of the infection time course, showing significant expression differences between different time-points, but no differential expression between compatible and incompatible interactions. In contrast, *RLK*, *PR1* and *PR10* genes revealed different expression profiles between the susceptibility and resistant coffee plant responses. These genes were up-regulated in an earlier time point in the incompatible interaction, and later on, although in a lower magnitude, in the compatible interaction. This could be related with defense responses resulting in the restricted fungal development in the incompatible interaction, while in the compatible interaction there is a delay in response and the infection progress is not affected. In addition, a putative correlation between the expression profiles of *Asp23673* and *PR* genes could support the possible involvement of *Asp23673* in degrading overaccumulated pathogenesis-related proteins (PR).

For the establishment of a VIGS system, a tobacco rattle virus (TRV) derived vector with an insert of the *PDS* gene was constructed and delivered to coffee plants by *Agrobacterium tumefaciens*. The same procedure was done for tobacco plants, as a control system. However, silencing of *PDS* using the developed VIGS system on both coffee and tobacco was unsuccessful, independently of the different experimental conditions tested, including different preparation of bacteria cultures, methods of inoculation and plant developmental stages.

Overall, our results confirmed previous studies and provided molecular insights on coffee-*H. vastatrix* interactions, and identified genes-of-interest for further functional studies. Additional studies are needed to develop and optimize a VIGS system in coffee plants, which would be a useful tool for better understanding the role of candidate resistance genes and contribute to improve the selection of coffee resistance to *H. vastatrix*, with the development of molecular markers.

Keywords: coffee leaf rust, compatible interaction, incompatible interaction, agroinoculation, tobacco

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AP – aspartic protease

Avr – avirulence gene

cDNA – complementary DNA

CIFC – centro de investigação das ferrugens do cafeeiro

CLR – coffee leaf rust

CoRSV – coffee ringspot virus

CT – threshold cycle

DNA – deoxyribonucleic acid

dsRNA – double stranded RNA

ETI – effector-triggered immunity

GADPH – glyceraldehyde 3-phosphate dehydrogenase

gDNA – genomic DNA

GFP – green fluorescente protein

GLP – germin-like proteins

HDT – Hibrido de Timor

HMC – haustorial mother cell

hpi - hours post inoculation

HR – hypersensitive reaction

ICO – International Coffee Organization

IM – induction medium

IPTG - Isopropyl β-D-1-thiogalactopyranoside

ISA – Instituto Superior de Agronomia

ITQB – instituto de tecnologia química e biológica

JA – jasmonic acid

LB – Luria-Bertani

LC-MS – Liquid chromatography-mass spectrometry

LM – light microscopy

LRR – Leucine-rich repeat

MCS – multiple cloning sites

MES - 2-(4 morpholino)-ethane sulfonic acid

mRNA – messenger RNA

PAL – Phenylalanine ammonia-lyase

PAMPs – pathogen-associated molecular patterns

PCR – polymerase chain reaction

PDS – phytoene desaturase

PR – pathogenesis-related

PRR- pattern recognition receptors

PTI – PAMP-triggered immunity

PTGS – post-transcriptional gene silencing

OxO – oxalate oxidase

qPCR – quantitative PCR

RISC – RNA-induced silencing complex

RLK – receptor-like kinase

RNA – Ribonucleic acid

ROS – reactive oxygen species

SA – salicylic acid

siRNA – small interfering RNA

SOD - superoxide dismutase

T-DNA – transfer DNA

TRV – tobacco rattle virus

U – enzyme unit

VIGS – virus induced gene silencing

1. INTRODUCTION

1.1. The Coffee plant

1.1.1. Taxonomy

According to Cronquist (1988)[1], coffee trees are classified as follows:

Kingdom: Plantae
Division: Magnoliopsida
Class: Magnoliopsida
Order: Rubiales
Family: Rubiaceae
Genus: *Coffea*

At the infrageneric level, Chevalier's (1947) classified the genus *Coffea* according to its geographical distribution and fruit characteristics into four sectors: *Paracoffea*, *Argocoffea*, *Eucoffea* and *Mascarocoffea*[2]. More recently, a new classification of two subgenera of *Coffea* was proposed, based on molecular, biochemical and cytogenetic data: *Coffea* subgen. *Coffea*, which comprises the majority of coffee species (95 species), including those used for producing the beverage coffee, and *Coffea* subgen. *Baracoffea* (9 species)[3–5]. Currently, the genus *Psilanthus* was brought together with *Coffea*, increasing the number of recognized species to 124[6]. However, only two species are used for commercial production: *C. arabica*, also known as Arabica Coffee and *C. canephora* also known as Robusta Coffee[5].

1.1.2. Origin and distribuition

Coffea subgenus *Coffea* is represented by 41 species in Africa, 58 in Madagascar, one in Grande Comore and three in the Macarenes, each region having 100% endemicity for its species[3].

C. arabica is mainly native to the highlands and relatively dry areas of Southwestern Ethiopia, having some populations arisen from South Sudan and North Kenya. This species was first domesticated in Yemen and then spread to other parts of Asia, Africa, and America[7–9]. Arabica coffee is grown throughout Latin America, in Central and East Africa, in India and to some extent in Indonesia[10].

C. canephora is native to the lowland forest of the Congo river extending up to Lake Victoria in Uganda. It is widely distributed in the tropical lowland forest of west and central Africa. Robusta coffee is grown in West and Central Africa, throughout South- East Asia and to some extent in Brazil, where it is known as Conillon[8,10,11].

C. arabica, contrary to *C. canephora*, is more adapted to cold environmental conditions and has center of origin on the Ethiopian Plateu. In the southwestern highlands of Ethiopia, in the remote mountain rainforest, some truly wild coffee populations can still be found[12]. *C. arabica* has 2 original botanical varieties: var. *typica* is presumed to have been originated from the native Ethiopian *C. arabica*, which was subsequently distributed into cultivation from the Yemen and var. *bourbon* arose also in Ethiopia and was cultivated by the French on Reunion Island. The majority of Arabica varietals is a product of these two varieties (obtained either by natural or artificial selection), e.g. "Caturra" is derived from var. *bourbon*[13].

1.1.3. General characteristics

Coffea spp. are evergreen, glossy-leaved shrubs or trees with 5-10m high and most adapted to a tropical forest habitat[14]. Their general botanical features include elliptical leaves, with pointed tips and

occurrence in pairs. Following the rainy season, flower clusters are produced in leaf axils and usually, a period of nine months spans between flowering and fruit ripening. The fruit starts as a two seeded green drupe, becoming red or yellow as it ripens[10]. One interesting aspect of this genus is the adaptation of its species to a wide range of environmental conditions. Arabica coffee originated in humid forests under shade conditions and its optimal temperature ranges from 18 to 21°C. However, it can also be cultivated in full sun due to its plasticity[15]. It is adapted to altitudes of 1300 to 1800 m and to a rainfall from 1600 to more than 2000 mm, with a dry season that lasts 3 to 4 months. In turn, an average temperature between 24 to 26°C, an altitude from the sea level up to 1200 m and rainfall superior to 2000 mm for 9-10 months are the ideal conditions for Robusta coffee[11,16]. For the consumers, *C. arabica* is preferred due to its flavor characteristics, quality and low-caffeine content compared to *C. canephora*, since this last one is bitter and has high-caffeine content.

1.1.4. Economic importance

Coffee is one of the world's favorite beverages with an estimation of 2.4 billion cups consumed per day worldwide[15]. Over 50 countries produce coffee in significant amounts, in many of these, earnings from coffee exports are of vital importance to the country's balance of payments. A further characteristic is that, with minimal exceptions, coffee is produced in developing countries. Consumption, on the other hand, takes place in industrialized countries and Brazil, which is the second largest coffee consuming country in the world, behind the USA. Total coffee production by all exporting countries was estimated by International Coffee Organization (ICO): 62.6 % of Arabica (A) and 37.4% of Robusta (R). Brazil was responsible for 36.7% (A/R) of total production followed by Vietnam (17.5%, R/A), Colombia (8.15%, A), Indonesia (6.1%, R/A), Ethiopia (4.46%, A) and Honduras (4.43%, A)[17].

Coffee production is however severely limited by major diseases, such as Coffee Leaf Rust (*Hemileia vastatrix*) and Coffee Berry Disease (*Colletotrichum kahawae*).

1.2. *Hemileia vastatrix* and coffee leaf rust

Coffee leaf rust (CLR) is a disease caused by the basidiomycete fungus *Hemileia vastatrix* Berkeley and Broome, from the order Pucciniales (rusts), and the major threat to Arabica coffee production worldwide. Currently, there are more than 50 known physiological races of this fungus[7,18].

1.2.1. Distribution and economic impact

The disease was first recorded in 1861 by a British explorer near Lake Victoria (East Africa) on wild *Coffea* species. The first great epidemic happened years later (1869) in Ceylon (now Sri Lanka), wiping out coffee cultivation from the greatest coffee growing land in the world at the time, with devastating social and economic consequences. This devastating outbreak was largely responsible for the switch in Sri Lanka's agricultural production from coffee to tea. Since then, coffee rust reached most coffee-growing countries worldwide, spreading through the coffee areas of Indian Ocean Basin and the Pacific between 1870 and 1920, to Africa Atlantic regions from 1950 to 1960, and to Latin America from 1970 to 1980[19]. Nowadays, *H. vastatrix* is endemic in all coffee growing countries, except Australia and Hawaii (CIFC's records). For this seriousness and major impact, CLR was ranked amongst the most catastrophic plant diseases of all times. Arabica coffee yield losses can reach up to 35% when the disease is not controlled[20], leading to losses of one to two billion US dollars annually, with a direct impact on the income and livelihood of thousands of farmers and labourers[21].

1.2.2. Symptoms

Coffee rust attacks the leaves and distinct symptoms can be seen, as yellow-orange lesions on the under surface of leaves. When just appearing, these are pale yellow spots of about 2 to 4mm in diameter, which increase in size and may coalesce with adjacent spots, reaching large areas of more than 5cm across. Soon after the appearance of the spots, clusters of uredospores are produced (sori), which form a yellow-orange dust-like deposit on the lesions (Fig. 1.1. A). The lesions increase in size, depending on the growth of the fungus inside the leaf. On a later stage of the disease, these lesions become necrotic, whereas the middle zone of the lesion continues to sporulate. CLR damages are caused due to the reduction in the photosynthetic leaf surface area. Severe attacks can cause heavy or even complete premature defoliation of the plant, premature fall of berries, dieback of the branches, and eventually slow death of the whole plant[13,22]. (Fig. 1.1. B). Disease severity varies according to humidity, genotype, crop load and nutritional status of the plants[23].

Figure1. 1 Coffee Leaf Rust symptoms: (A) Pustules of uredosporic sori on the abaxial side of a coffee leaf, adapted from Talhinhas et al. (2017)[7]; (B) Severe defoliation of coffee plants due to the disease, photo by Vitor Várzea.

1.2.3. Pathogen life-style, life-cycle and infection

H. vastarix is an obligate biotroph that depends entirely on living host tissues for its growth and reproduction. It is a hemicyclic fungus producing uredospores, teleutospores and basidiospores. Uredospores and teleutospores are produced in the same sorus, but at different times. Uredospores are dikaryotic and represent the asexual cycle, re-infecting the leaves whenever environmental conditions are favorable. Teleutospores differentiate into basidiospores (sexual phase), which have been rarely found and seem to be unable to infect the coffee leaves. Also, no alternate host plant that can be infected by basidiospores was identified. The function of these two spore stages remains unclear in the life cycle of *H. vastatrix*[15,24].

When the ideal conditions for the infection process are found, usually 22°C to 24°C as the optimum temperature for germination of the fungus[15] and the presence of water, the following events occur: adhesion of the uredospore to the host leaf surface, and upon recognition of chemical and physical properties, uredospore germination and appressorium differentiation over stomata takes place, from which penetration hyphae is formed and grows through the substomatal chamber. This structure ramifies and produces two thick lateral branches at the tip, resembling an anchor. From each lateral branch of the anchor, a hypha called haustorial mother cell (HMC) is born, giving rise to haustoria (specialized fungal feeding structures) that starts by invading the subsidiary cells. More intercellular hyphae and haustoria are formed in the cells of the spongy and the palisade parenchyma and upper epidermis. About 21 days after inoculation, a dense mycelium and an uredosporic sorus over the stomata are visible[24] (Fig. 1.2.).

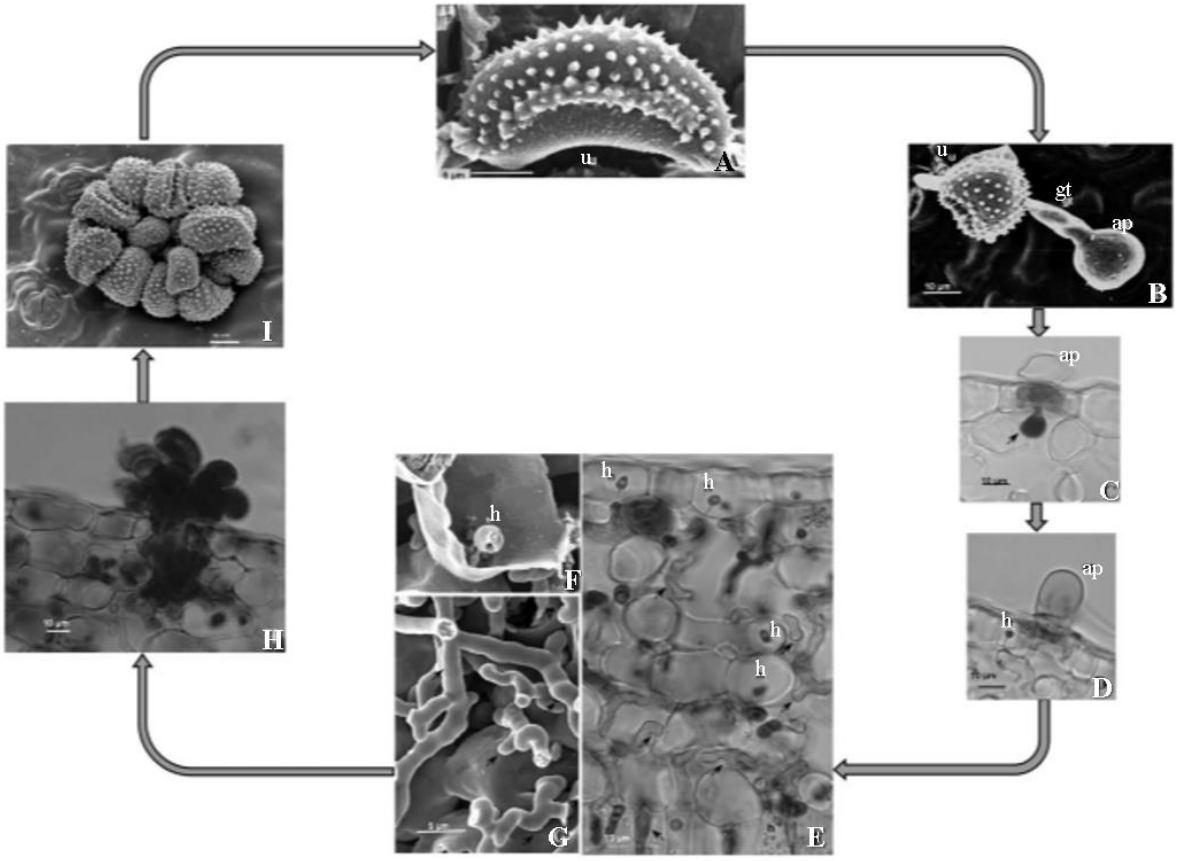


Figure1. 2 *Hemileia vastatrix* infection process. (A) Uredospore (u), scanning electron microscopy (SEM). (B) Germinated uredospore (u) with germ tube (gt) and appressorium (ap) over stomata on the lower surface of the coffee leaf, 17hours post inoculation (hpi), SEM. (C) Appressorium (ap) over stomata and penetration hypha (arrow), 24 hpi, light microscopy (LM). (D) Appressorium (ap) over stomata and intercellular hypha with an haustorium (h) within a subsidiary cell, 48 hpi, LM. (E) Intercellular hyphae (arrows) and haustoria (h) within epidermal and mesophyll cells, 20 days post inoculation (dph), LM. (F) Haustorium (h) within a spongy parenchyma cell, 20 dph, SEM. (G) Intercellular hyphae (arrows) in the spongy parenchyma, 20 dph, SEM. (H) Uredosporic sorum protruding through the stomata in a bouquet shape, 21 dpi, LM. (I) Uredosporic sorum, 21 dph,SEM. Adapted from Talhinhas et al. (2017)[7].

1.2.4. Disease control

Chemical control of CLR is the most largely used method to contain the disease. Preventative treatments are typically carried out with copper-based fungicides, whereas curative treatments are conducted with systemic fungicides (e.g. epoxiconazole, pyraclostrobin), but a combination of both is also used[7,13]. However, this is extremely expensive (chemicals, manpower, water availability, machinery and equipment, etc.), represents an environmental hazard, can be difficult to apply and requires technical inputs which are not available in developing countries. For instance, in most Latin America countries, chemical spraying is unsuitable due to the steep slopes of the coffee plantations[7].In alternative, breeding of coffee plants for rust resistance is considered to be the best disease management strategy, both environmentally and economically[24]. Since 1955, the Centro de Investigação das Ferrugens do Cafeeiro (CIFC)/Instituto Superior de Agronomia (ISA) has played a major role in supporting breeding programs in coffee growing countries, by developing coffee genotypes manifesting a wide spectrum of resistance to fungal races from different parts of the world, which originated over 90% of the resistant coffee varieties currently cultivated worldwide. One of the major breakthroughs was the discovery of Hibrido de Timor (HDT), a natural hybrid between *C. arabica* and *C. canephora*, with the majority of its derived populations offering resistance to all rust races known at the time. Some further developed lineages from these original populations were indeed

the resistance sources of the majority of currently grown rust-resistant varieties. However, in some countries, these improved varieties are gradually losing their resistance to rust since *H. vastatrix*'s virulence is increasing[24].

1.3. Plant defenses and coffee resistance to rust

Plants employ a multi-layered system of defenses to protect themselves from various pathogens. In addition to passive mechanisms, plants have evolved at least two lines of active defenses[25–28]. The first line provides basal defense against all potential pathogens and is based on the recognition of conserved pathogen-associated molecular patterns (PAMPs), by so-called pattern recognition receptors (PRRs) that activate PAMP-triggered immunity (PTI). The second layer of defense is activated when a given pathogen-derived molecule, called effector, is 'specifically recognized' by plant receptor proteins encoded by R genes, resulting in effector-triggered immunity (ETI). Once activated, both PTI and ETI induce a downstream of similar responses, such as rapid accumulation of reactive oxygen species (ROS), changes in cellular ion fluxes, activation of protein kinase cascades, production of stress-related hormones, cell wall modifications and changes in protein and gene expression[27,29]. ETI, when compared with PTI, is associated with more sustained and robust immune responses including hypersensitive reaction (HR), a form of programmed plant cell death localized at the infection sites[27] which is considered to be one of the most important factors in the restriction of the pathogen growth, particularly of obligate biotrophs[30,31].

Coffee resistance to rust is a race-specific resistance, which is governed by the gene-for-gene model, stating that for each host resistance gene (R) there is a corresponding pathogen avirulence gene (*avr*)[24], and specific interactions of their gene products leads to a response of resistance or susceptibility. Coffee resistance is characterized by restricted fungal growth (pre- or post haustorial resistance) associated with several plant responses, with the increase in the activity of phenylalanine ammonia-lyase (PAL), β-1,3-glucanase, chitinases and oxidative enzymes (e.g. peroxidases and lipoxygenases)[24,32,33]. Recent proteomic studies, which generated a global protein profile of incompatible and compatible *C. arabica-H. vastatrix* interactions, identified in the apoplastic fluid defense-related proteins, such as proteases, hydrolases and oxidases[34]. More specifically, germin-like (GL) proteins and aspartic proteases nepenthesin-2 were identified, showing a higher expression in the resistant than in the susceptible samples[35]. GL proteins are involved in various enzymatic activities, including superoxide dismutase (SOD) and oxalate oxidase (OxO), also acting as receptors in signal transduction processes[36]. Aspartic proteases nepenthesin-2 have been associated to protein catabolism, folding, sorting and degradation[35]. Several genes putatively involved in signalling, recognition and defense were also identified in coffee-*H. vastatrix* interactions, including subtilisin-like protease, *WRKY* transcription factors, glucosyltransferases, lipoxygenases and phenylalanine ammonia-lyase. Also, pathogenesis-related genes (PR), namely *PR1* and *PR10*, and receptor-like kinase (*RLK*) genes are suggested to have an involvement in the resistance response in this interaction, having a stronger expression in host cells in incompatible interactions. PR proteins represent a protein group strongly induced by pathogen attack, which comprise 17 families. In particular, PR1's function is still unknown, although its synthesis seems to be induced by salicylic acid (SA). LC-MS studies of different phytohormones suggested the involvement of a salicylic acid-dependent pathway in coffee resistance to *H. vastatrix*, which is in accordance with gene expression data[37]. On the other hand, *PR10*, acting as a ribonuclease, seems to be related with the jasmonic acid (JA)-dependent resistance pathways. For instance, in the interaction of coffee with *Colletotrichum kahawae*, the involvement of JA was shown to be more relevant for the outcome of resistance than SA[38]. *RLK* are transmembrane-receptors that perceive external and internal signaling molecules, that once activated form complexes with other receptors resulting in distinct signaling pathways which transmit the signal into the cell[39]. In

the interaction coffee-*H. vastatrix*, an activation of this gene was detected throughout the infection process at a higher magnitude in the incompatible interaction[37,40–43]. In coffee plants susceptible to *H. vastatrix*, during later stages of the infection, an increase in the abundance of defense-related proteins has also been detected[32,34,44], suggesting a delay in response. Furthermore, HR of guard and subsidiary cells and the encasement of some haustoria have also been observed, but in a low percentage of infection sites. Nevertheless, these plant defense responses occur too late to efficiently prevent fungal growth and sporulation[24,44]. A better understanding of genes and defense pathways involved in the resistance of coffee to rust is crucial to identify sources of more durable resistance and increase the life time of commercially resistant cultivars.

1.4. Functional gene analysis

Upon identification of candidate genes potentially involved in resistance, it is necessary to validate and better elucidate their role through functional studies. Linking each gene with a specific phenotype or biological function is indeed what matters as the basic information necessary for any improvement breeding program. As a first-step analysis, the transcription levels and differential expression of the candidate genes can be quantified and studied to understand their function using quantitative Real Time PCR. Further functional genomics approaches usually involve experiments of stable plant transformation and transgenic line characterization to confirm gene function. However, this process can be simplified by the use of other approaches, such as virus induced gene silencing (VIGS), which allows to directly altering expression of the gene sequence of interest and subsequently identifying the mutant phenotype produced after changing their expression. This is particularly important for woody species such as coffee, for which genetic transformation is far from a routine procedure[8].

1.4.1. Quantitative polymerase chain reaction (qPCR)

Quantitative polymerase chain reaction (qPCR) is a method that uses fluorescent technology to precisely follow the amplification of the DNA during the PCR reaction[45]. qPCR assays can be based on DNA-binding dyes, like SYBR Green. When it is used, a sequence detector reads the increase of fluorescence emission in ‘real time’ observed during the polymerization step, indicating the target amplification during the reaction[46]. Consequently, it is possible to monitor the increasing amount of amplified DNA[45]. Basically, the PCR cycle connected to the product exponential growth is tightly associated with the quantity of the initial cDNA template, providing an estimation for the level of mRNA expression in the tissues. A parallel measurement of reference genes is needed to adjust for intrinsic experimental variations and accurately normalize the data. The chosen reference genes should be genes which are essential for the maintenance of cellular function, and their transcription is not affected within the experimental context under investigation[47].

1.4.2. Virus-induced gene silencing

VIGS consists of a transient transformation strategy that can be used as a powerful and rapid tool in host gene validation for loss-of-function[48]. It employs recombinant viruses to specifically reduce host endogenous gene activity through plant innate silencing mechanisms called Post-Transcriptional Gene Silencing (PTGS)[48]. During viral replication, double stranded RNA (dsRNA) is produced by an RNA-dependent RNA polymerase, which triggers PTGS. Dicer-like enzymes cleave those dsRNAs producing siRNAs (duplexes of 21-24 nucleotides in length). This siRNA binds to and activates the RNA-induced silencing complex (RISC), as it guides and cleaves the viral RNA in a homology-dependent manner. This information is spread systemically throughout the plant. When a viral vector

carries a host cDNA fragment of interest, virus infection triggers degradation of the cognate plant mRNA through a homology-dependent RNA degradation mechanism that results in a down-regulation of the target gene and thus in a loss-of-function phenotype of the host[49–51](Fig. 1.3).

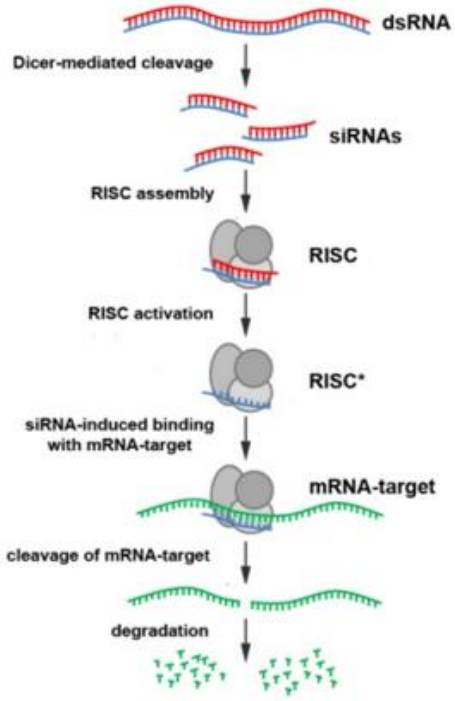


Figure1. 3 - Mechanism of action of interference RNA. dsRNA are cleaved by Dicer enzymes, producing siRNAs. These structures bind to RNA-induced silencing complex, guiding the complex to a mRNA target, leading to its degradation. Adapted from Petrova et al. (2013)[51].

Many VIGS vectors have been used to study gene function in plants, being the Tobacco Rattle Virus (TRV) the most popular to facilitate the silencing of the target endogenous gene because of the mild symptoms associated with its infection and its wide range of potential plant hosts. TRV-based VIGS consists on a binary system modified for transient expression mediated by the plant's pathogen *Agrobacterium tumefaciens*: TRV1, which encode proteins needed for replication and movement within the host plant, and TRV2 which has genes encoding the coat protein, viral replication, and nonstructural proteins unessential for plant infection. Therefore, to construct the vector used in VIGS, these nonstructural protein-encoding genes in TRV2 are replaced with multiple cloning sites (MCS), making it possible to insert fragments of the target gene desired to be silenced. Then, both TRV1 and TRV2- derived constructs are then delivered into the plant through an *A. tumefaciens* mediated manner[48,51,52].

Phytoene desaturase (*PDS*) is used as a reporter gene in VIGS to control the silencing procedure. It encodes an enzyme involved in carotenoid biosynthesis pathway and its silencing leads to a photobleaching phenotype[52]. In Figure 1.4. it is possible to see a control plant that was infiltrated with a vector harboring a fragment of the green fluorescent protein (*GFP*) reporter gene (A), and a plant with white chlorosis caused by the silencing of *PDS* by VIGS (B).

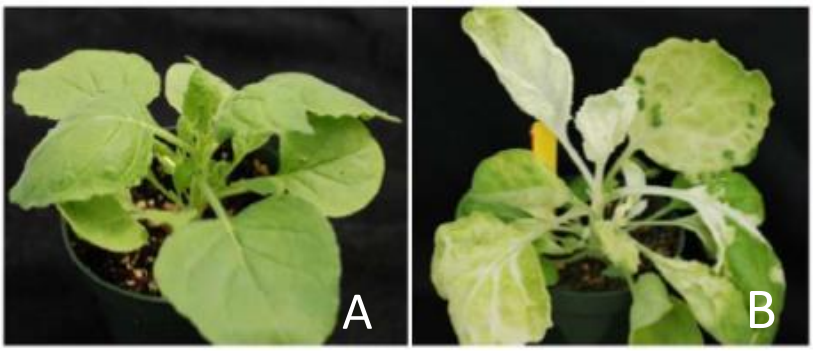


Figure1. 4 - VIGS with reporter gene on tobacco plants. (A) Control plant with TRV2::*GFP*; (B) plant withTRV2::*PDS* showing induced chlorosis. Adapted from Senthil-Kumar&Mysore (2014)[50].

Virus-induced gene silencing has been widely used for functional gene analysis in a broad range of plants, such as tomato, tobacco, *Arabidopsis*, rose, cotton, apple and *Stiga hermonthica*[53], although it has a better efficiency in some species than in others[54]. In tobacco, TRV-VIGS can be accomplished within 4 weeks, and the stably silenced plants can be maintained for more than 2 years for performing various analyses. VIGS is a quick and easy powerful tool for knocking-down genes in plants that does not alter the plant genome[53]. It can overcome functional redundancy of gene families and avoids genotype-specific effects between different genetic backgrounds. However, silencing is never 100% efficient. Thus, care must be taken when interpreting the results to better understand where the process could be optimized. Successful VIGS depends upon the virus vector construction, inoculation techniques and environmental conditions[55,56]. Taken together, these abilities make TRV-VIGS one of the best methods for both forward and reverse screens for high-throughput gene function analyses.

1.5. Objectives

In this research the focus was to study and validate the functional role of a group of genes previously shown in other studies to be putatively involved in resistance mechanisms of coffee to rust.

Specifically, this work aimed at the:

1 – Identification and characterization of gene expression profiles of selected genes in coffee plants exhibiting differential behavior (susceptibility and resistance) towards *H. vastatrix* infection:

> Arabica coffee var. Caturra was inoculated with two *H. vastatrix* isolates establishing an incompatible and compatible interaction. Light microscopy was used to analyze the differential development of the fungus in both type of responses (resistance and susceptibility) and validate the time-points selected for gene expression analysis. Afterwards, expression analysis of 7 genes (*GL18058*, *GL22853*, *Asp23673*, *PR1*, *PR10* and *RLK*) involved in recognition, signaling and defense mechanisms was performed by qPCR, to study differential patterns of gene expression in the two contrasting interactions.

2- Development and optimization of a VIGS system in coffee plants, to further validate the role of candidate genes involved in defense pathways of coffee to *H. vastatrix*:

> A tobacco rattle virus (TRV) derived vector carrying a fragment of the coffee *PDS* gene was constructed and later delivered into coffee plants, using *Agrobacterium tumefaciens*. For comparative analysis using a control system, the same procedure was developed for tobacco. In this work, a comprehensive set of experimental conditions were tested, including different protocols of preparation of bacteria cultures, methods of inoculation and plant developmental

stages. Infiltration results were assessed by visual inspection of photobleaching symptoms and PCR detection of TRV coat protein gene.

2. MATERIALS AND METHODS

2.1. Gene expression studies

2.1.1. Plant material and inoculation

Coffea arabica L. plants of variety Caturra, from Angola (CIFC 19/1, genotype S_H5) were used to establish both compatible and incompatible interactions. Young plants were inoculated with fresh uredospores of *H. vastatrix* isolate 71 from Mozambique (race VI, genotype $v_?$) for establishing an incompatible interaction and isolate 1427 from Kenya (race II, genotype v_5) for establishing a compatible interaction, thus providing a resistance and susceptible response, respectively. For each interaction, 3 pairs of leaves were inoculated in each of 3 replicate plants, and 3 pairs of leaves were mock-inoculated (sprayed with water) in control plants.

Following a routine procedure used at CIFC, fresh uredospores of *H. vastatrix* (1 mg per pair of leaves) were placed with a scalpel on the lower surface of young leaves and then brushed gently with a camel-hair brush (D'Oliveira 1954-57) (Fig. 2.1 A, B). The inoculated leaves were sprayed with distilled water (Fig. 2.1 C) and both inoculated and control plants were placed for 24 hours under darkness in moist chambers, after which they were moved to greenhouse benches.

Leaves from resistant, susceptible and control plants, were collected at 24, 48, 72 and 96 hours post inoculation (hpi). The central vein of each leaf was removed using a blade and immediately frozen in liquid nitrogen and conserved at -80°C until used.

Figure 2. 1 - Inoculation of coffee leaves with *H. vastatrix* uredospores. (A) Placing uredospores with a scalpel, (B) brushing uredospores with a brush, (C) spraying inoculated leaves with distilled water.

2.1.2. Light microscopy observations

Germination "*in vivo*" and appressoria formation were evaluated on leaf pieces (±5cm^2), 24hours after the inoculation, following the technique of Silva et al. (1999)[57]. Briefly, leaf pieces were painted with transparent nail polish on the lower surface. About 24h later, the nail polish (leaf replica) was removed with tweezers, stained and mounted with blue lactophenol. Counts of the germinated uredospores and appressoria formed on stomata were made on a minimum of six microscope fields of 100 uredospores each/experiment.

To evaluate fungal post-penetration stages, infected leaf cross-sections (20-25 µm), made with a freezing microtome (Leica CM1850) were stained and mounted in cotton blue lactophenol[57]. Data were recorded from 60 infection sites/experiment along the time-course, 24, 48, 72 and 96 hpi. Observations were made using light microscopes (Leitz Dialux 20 and Leica DM-2500).

2.1.3. RNA extraction and cDNA synthesis

Flash-freezed leaf samples were thoroughly ground to powder with a mortar and pestle stored at -80°C. Total RNA was extracted using *Spectrum Plant Total RNA kit (Sigma)*, following the manufacturer's instructions. RNA concentration and purity of each sample was measured at 260/280 nm and 260/230 nm using a spectrophotometer (ND-1000, Thermo Scientific). RNA integrity was verified by electrophoresis in 1% (w/v) agarose gel, loading 1 µl (with about 200 ng/ µl) of RNA.
RNA samples to be used in qPCR experiments were treated with TURBO DNA-*free* Kit (Thermo Scientific) to eliminate residual DNA. For this procedure the following steps were performed: 35-40 µl of RNA extracted volume, 1 µl of Turbo DNAse (2U), 5 µl of 10x turbo DNAse buffer (0.1 V) and DEPC water to a final volume of 50 µl were mixed in a tube and incubated at 37°C for 1 hour. Then, 5 µl of DNAse Inactivation Reagent (0.1V) was added and the mixture was left 5 minutes at room temperature. Samples were then centrifuged for 2 min at 10000 rpm, and the supernatant was transferred to a new tube. For precipitation of RNA: 1/10 NaAc and 2.5 vol of cold 100% ethanol were added to each sample. Samples were left at -80°C for 45 min and then centrifuged at 14000 rpm, 4°C, for 30 min. Supernatant was discarded and the pellet was washed twice with 750 µl of 70% ethanol. Each sample was then centrifuged at 14000 rpm, 4°C for 10 min. The pellet was resuspended in 14 µl of H_2O DEPC. RNA concentration/purity and quality were estimated as explained above.

Presence of genomic DNA (gDNA) in RNA samples leads to a misleading interpretation of the gene expression, since DNA is also amplified. To check for DNA contamination, qPCR analysis was performed on the crude RNA samples using 1 µl, as described in section 2.1.5 of this work.

Complementary DNA (cDNA) was synthesized using the *Omniscript RT kit (Qiagen)*, from 1000 ng of total RNA. For each sample, 2 µl of oligo-dT (10 µM), 2µl of dNTP mix (5mM), 1.0 µl of Buffer RT (1X), 1 µl of RNase Ribolock inhibitor, 1 µl of Omniscript Reverse Transcriptase (10 U/µl) and RNase-free water to a final volume of 20 µl, were mixed and briefly centrifuged. The reaction was incubated at 37°C for 1 hour. All cDNA samples were stored at -20°C.

2.1.4. Selection of candidate genes for coffee resistance and primer design

Genes of interest were selected according to their putative activity related to plant defense mechanisms (Table 2.1.). With the exception of *LRR830.5*, *LRR1241* and *β-Glu 19724* genes, the remaining 8 genes were previously described in coffee[35,37,38,58,59]. Sequences of genes *LRR830.5* (predicted putative LRR receptor-like serine/threonine-protein kinase At2g16250), *LRR1241* (predicted LRR receptor-like serine/threonine-protein kinase At4g08850) and *β-Glu 19724* (predicted β-glucosidase 24) were retrieved from a coffee RNA-seq database (Bioproject database, NCBI, accession number PRJNA271934). Specific primers were designed with PrimerSelect version 5.0 (DNAStar, Inc, USA) using the following parameters: amplicon length 70 and 200 bp; size between 17 and 22 bp; annealing temperature (Ta) between 58 and 62°C and GC content ± 50%.
Ubiquitin and *GADPH* were used as reference genes (Table 2.1.) as established in previous studies[37].

2.1.5. Quantitative real-time PCR

All qPCR experiments were carried out in 96-well *Multiplate (Bio-Rad)* white plates sealed with *Microseal 'B' Seals (Bio-Rad)*, using *SYBR Green Supermix (BioRad)* in an *iQ5 realtime thermal cycler (Bio-Rad, USA)*. Each 25 µl reaction was composed of 12.5 µl of SYBR Green Supermix (Bio-Rad USA), 0.5 µl of each primer (0.2 µM), except for *Asp23753*, used at a concentration of 0.1 µM, and *PR1* and *RLK*, both used at 0.15 µM, 4 µl of cDNA template at a concentration of 50-200 ng/µl, and the rest of sterile distilled water to make the total volume.

Table 2. 1 – Primers sequences for qPCR analyses of target and reference genes. *PCR efficiency when cDNA was at a concentration of 50ng/µl. ** PCR efficiency when cDNA was at a concentration of 200ng/µl. Ref- Reference, TS – This study

Gene	Description (Acession number)	Primer sequence (5'-3')	Amplicon length (bp)	Ta (°C)	PCR efficiency (%)	Ref
GL18058	Germin-like protein, (XM_027248956.1)	Fw: CCGCACTATCATCCAAGAGC Rv: GAAAGACCATCACCTCACCCT	127	60	93	58
GL22853	Auxin-binding protein ABP20-like, (XM_027210983.1)	Fw: GTTCCCTGGTGTGAATGGTC Rv: TCTCCAATGAAGAAATGAAGCCA	152	60	96	58
Asp23673	Probable Aspartyl protease, (XM_027227257.1)	Fw: CATCCAACACCAAGAAGCCA Rv: CTCATCCAGACAACCAGCAG	146	60	82	58
Asp203	Aspartic protease in guard cell1-like, (XM_027267645.1)	Fw: CTTTCTCCTACTGCCTCGTG Rv: TTCGCCTATTTGATAAGCCGAT	200	60	102	58
PR10	Pathogenesis-related 10, (CF589103)	Fw: GCCACCATCCTTGAAGAGAA Rv: CAACTCTCTGCTTGGCAGTCT	151	55	83	59
PR1	Pathogenesis-related 1, (Scaffold170607)	Fw: GCCCGTAAAGTCACCTGT Rv: AACTACGCTGCCAAAATCC	105	62	94	38
PR1b	Pathogenesis-related 1b, (DQ335594)	Fw: GATTACCTGGACGCCCATAA Rv: GCTGCCAGTTTTCTCCATA	170	-	-	59
RLK	Putative Receptor-Like Kinase, (CF589181)	Fw: ATGGGAGAAAAGAATGGCAGAAG Rv: GGCCAATTACAGTTTGAAAACACC	189	62	86	59
LRR 830.5	Probable LRR receptor-like serine/threonine-protein kinase, (XM_027227582.1)	Fw: AAGAGACGAATGGGAGTAGG Rv: GCCATTCAATACATGCTTCAG	179	-	-	TS
LRR 1241	Predicted LRR receptor-like serine/threonine-protein kinase, (XM_027261095.1)	Fw: AAACTTCACAACATGCCTGAG Rv: TGAGAATTTGAGCAGAAGCCA	119	-	-	TS
β-*Glu* 19724	Predicted β-glucosidase 24, (XM_027217201.1)	Fw: GTTATACAGCTCGTTTTGGGA Rv: GAGGATCTTCTTCTTAGGGTT	188	-	-	TS
		Fw: GACTCAAACCGTAGGTCCTG Rv: GACCTTCTGAGACAGTAAGACC	194	-	-	TS
GADPH	Glyceraldehede-3-phosphate dehydrogenase, (SGN-U619744)	Fw: TTGAAGGGCGGTGCAAA Rv: AACTAGGGTGCATCCTTGCT	95	60	106*; 103**	37
Ubi	Ubiquitin-like protein, (AF297089)	Fw: AACATTGAGGGTGGTTCTGTT Rv: GCAGAAAACCAACTAAGACCTAACAA	80	60	95*; 105**	37

The following thermal cycling conditions were used: 3 min at 95°C for polymerase activation denaturation for 45 cycles at 95°C for 10 seconds, and annealing at the respective temperature for each gene (Table 2.1) for 30s. A melting curve analysis was performed at the end of the PCR run over the

range 55-95°C, increasing the temperature in a stepwise fashion by 0.5°C every 10s. Each set of reactions included a negative control with no template. Dissociation curves and agarose electrophoresis were used to analyze non-specific PCR products. LinRegPCR version 2012.0 software was used to determine the primer's amplification efficiencies for each gene of interest. Three biological replicates and two technical replicates were used for each sample, and only technical replicates with a sample variance of CTs below 0.5 were considered for the experimental analyzes. Relative gene expression (fold change) was calculated according to Helleman et al. (2007). Dissociation curves and agarose gel electrophoresis were used to analyze non-specific PCR products.

2.1.6. Data analysis

For statistical analysis of cytological data, Student's t test was applied using IBM1SPSS1 Statistics version 20.0 (SPSS Inc., USA) software and arcsine-transformed percentages were used where appropriate. Statistical significance ($p \leq 0.05$) of gene expression between compatible and incompatible coffee interactions, and between different time-points, was determined by the non-parametric Mann–Whitney U test using IBM1SPSS1 Statistics version 21 (SPSS Inc., USA) software.

2.2. Virus-induced gene silencing (VIGS)

2.2.1. Plant material

For the VIGS experiments, *Coffea arabica*'s hypocotyls (8 week-old), seedlings (10-12 week-old) and grown plants from var. Caturra kept at CIFC's greenhouses at a temperature between 16°C and 27°C (average minimum and maximum temperatures respectively) were used. In parallel, for control experiments to evaluate VIGS efficiency, *Nicotiana rustica*'s seeds were sown in seedbeds in CIFC's greenhouse conditions, during 3-4 weeks.

2.2.2. Isolation of PDS genes from coffee and tobacco

The coding sequences (CDS) of *PDS* (Phytoene desaturase) genes of coffee and tobacco were obtained from GenBank (DQ357179.1 and DQ469932.1, respectively) and aligned to perform a homology analysis. Primers were designed with a degenerated base, based on these sequences (Supplementary Fig. 1):

F: ATGCAGAACCTGTTTGGAGAA;
R: GTTMAGTGCCTTTGACATGGC, (M = A/C).

The collection and grind of coffee and tobacco plant material and total coffee RNA extraction were performed as previously described in section 2.1.3. For *N. rustica*, total RNA was extracted with *RNeasy Plant Mini kit (Quiagen)*, according to the manufacturer's instructions. cDNA was synthesized as explained in section 2.1.3., using 2000 ng of RNA.

Amplification of the *PDS* gene fragments from coffee (*CaPDS*) and tobacco (*NrPDS*) was performed under the following conditions: 20 µl reaction mixtures were prepared with 4 µl of GoTaq Flexi buffer (1X; Promega), 2 µl of $MgCl_2$ (2.5 mM; Promega), 0.16 µl dNTPs mix (0.08mM; Promega), 0.4 µl of each forward and reverse primers (0.2 mM), 0.1 µl of GoTaq G2 Flexi DNA Polymerase (5U/ µl; Promega), 4 µl of cDNA template (50 ng/µl) and water to the final volume. A reaction with 4 µl of water instead of cDNA, was used as a negative control. PCR was carried out in *MyCycler Thermal Cycler System* (*Bio-RAD)* with an initial denaturation step at 95°C for 4 min, followed by 40 cycles of denaturation at 95°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1 min, and one final step of extension at 72°C for 10 min. A 4 µl aliquot of each PCR reaction was then directly run on a 1% (w/v) agarose gel in 1X TBE buffer at 80 V for 45 min. After gel electrophoresis, PCR products were visualized using *Gel Documentation System Omnidoc (Cleaver Scientific)*. The desired bands of both *CaPDS* and *NrPDS* were extracted from the gel using a scalpel blade and purified using

the *Wizard SV Gel and PCR Clean-Up System kit (Promega)*, according to the manufacturer's instructions. Both purified products were sent to STAB VIDA to be sequenced.

2.2.3. Cloning of *CaPDS* and *NrPDS* fragment into pGEM-Teasy vector and *E. coli* transformation

The *pGEM-T Easy Vector Systems (Promega)* kit was used. The purified *CaPDS* and *NrPDS* fragments were individually ligated to pGEM-T Easy vector according to the manufacturer's instructions using17,25 ng of the insert for a 3:1 (insert:vector) molar ratio. These ligation reactions were left to incubate overnight at 4°C. Afterwards, transformation was performed, according to the manufacturer's instructions using JM109 High Efficiency Competent Cells (*Promega*), and 100µl of each transformation culture were plated onto Luria-Bertani (LB) culture medium supplemented with 100µg/ml ampicillin, 0.5mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and 80µg/ml X-Gal plates and incubated overnight at 37°C. The presence of IPTG, a compound that induces protein expression, and X-gal, a compound hydrolyzed by β-galactosidase, allows the identification of the recombinant bacteria, since the inserted fragment interrupts the coding region of β-galactosidase. Therefore, if the insert is present on the vector, X-Gal is not hydrolyzed and colonies will be white. If the vector remains intact, colonies will turn blue.

After incubation, white colonies were selected and a colony PCR was performed, using the universal primers SP6 and T7, under the following conditions: 20 µl reaction mixtures were prepared with 4 µl of GoTaq Flexi buffer (1X; Promega), 2 µl of $MgCl_2$ (2.5 mM; Promega), 0.4 µl dNTPs mix (0.5mM; Promega), 0.8 µl of each forward and reverse primers (0.4 mM), 0.2 µl of GoTaq G2 Flexi DNA Polymerase (5U/ µl; Promega), 2 µl of water where an isolated colony picked with a toothpick was suspended, and water to the final volume. PCR was carried out with an initial denaturation step at 95°C for 1 min, followed by 35 cycles of denaturation at 95°C for 10 s, annealing at 50°C for 30 s and extension at 72°C for 30 s, and one final step of extension at 72°C for 5 min. PCR products were run on an agarose gel and visualized as described in section 2.2.2.

PCR positive colonies were grown in 5 ml of LB medium containing 100 µg/ml ampicillin overnight at 37°C on a gyratory shaker at 150 rpm. Minipreps were performed using the *QIAprep Spin Miniprep Kit (Quiagen):* The bacterial overnight culture was pelleted by centrifugation for 6 min at 13000 rpm at room temperature, using only 2 ml each time in independent centrifugations. The remaining protocol was performed following the manufacturer's instructions. The concentration of the extracted plasmid was measured in ND-1000 and sent to sequencing at Genewiz. Additionally, the presence of the *PDS* insert in the plasmid samples was confirmed by PCR under the same conditions used for colony PCR, with an annealing temperature of 56°C.

2.2.4. Viral plasmid construction

TRV1 and TRV2 VIGS vectors described by Liu et al. (2002) were kindly provided by Prof. Danielle Barros, Universidade Federal de Pelotas, Rio Grande do Sul, Brazil. To prepare TRV2::PDS vector (Supplementary Fig. 2), *CaPDS* and *NrPDS* inserts were extracted from the pGEM-T Easy vector by digestion with *FastDigest EcoRI (Thermo Scientific).* The same enzyme was used to digest the empty TRV2 vector along with *FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific)*, to prevent the recircularization of the vector.

For both coffee and tobacco, a total of 2000 ng of the extracted plasmid was digested using 1µl of enzyme in 1X of the specific buffer, in a 20 µl reaction, at 37°C for 30 min, followed by 5 min of inactivation of the enzyme at 80°C. A total of 1000 ng of the TRV2 vector was also digested with 1 µl of FastDigest EcoRI and dephosphorylated using 1 µl of phosphatase (1U) in 1x FastDigest Buffer at 37°C for 30 min, followed by 20 min of inactivation of both enzymes at 80°C. Electrophoresis of the

digestions' products was performed on a 1.2% agarose gel at 80V. The expected bands (345 bp corresponding to *PDS* and 9663 bp corresponding to linearized TRV2) were extracted and purified using *High Pure PCR Product Purification Kit (Sigma)*, following the manufacturer's instructions.

After the purification of the bands, ligation was performed using two molar ratios, 1:1 and 3:1 (insert:vector).
For the 1:1, 100 ng of the TRV2 and 5,6 ng of *PDS* were used.
For the 3:1, 100 ng of the vector and 16,8 ng of *PDS* were used.
Vector and insert were then ligated using 5U of T4 DNA ligase (*Rapid DNA Ligation Kit, Thermo Scientific*), in 1X reaction buffer at 22°C for 5 min and stored at 4°C until used for transformation.

2.2.5. Transformation of *Agrobacterium tumefaciens* with TRV2::PDS vector

Cells from the *A. tumefaciens* strain C58C1 were turned competent for further transformation. Bacterial cultures were grown on LB plates, supplemented with 50 mg/ml rifampicin and 50 mg/ml gentamicin, at 28°C during 3 days. After incubation, individual colonies were selected and picked with sterile toothpicks and left to grow on 5 ml LB medium, with the same antibiotics, overnight at 28°C on a gyratory shaker at 180-200 rpm. On the next day, 2 ml of that culture were added to 50 ml of LB medium and incubated under the same conditions until the culture grew to an OD_{600} of 0.7. The culture was chilled on ice, and centrifuged at 3000 g for 5 min at 4°C. The pellet was resuspended in 1 ml of 20 mM $CaCl_2$ solution. Bacterial cultures were distributed on 1.5 ml *eppendorf* tubes. Some of these tubes were frozen in liquid nitrogen and stored at -80°C, after adding glycerol 50% at 1:1, the rest was directly used for transformation.

A. tumefaciens transformation was done by adding 1 µg of the plasmid to the frozen cells, which were then thawed in a water bath at 37°C for 5 min and placed on ice for 30 min. Cultures were later spread on prewarmed LB plates containing the same antibiotics previously used for bacteria growth plus 50 mg/ml kanamycin, used for selection of colonies with the TRV2::PDS vector. Colonies were observed three days after incubation at 28°C and a colony PCR was performed to confirm positive colonies, using both *PDS* and TRV2 coat protein primers (F: CTGGGTTACTAGCGGCACTGAATA; R: TCCACCAAACTTAATCCCGAATAC). Plasmid DNA was extracted using *QIAprep Spin Miniprep Kit (Quiagen)* with minor modifications. A single colony was inoculated into 20 ml liquid LB medium and grown for 3 days at 28°C, on a gyratory shaker at 200 rpm. The bacterial culture was pelleted by centrifugation for 5 min at 3600 rpm at room temperature, using only 2 ml of the culture each time in independent centrifugations. The remaining procedure was followed according to the manufacturer's instructions.

Since the protocol above gave a low yield of extracted plasmid, the *Wizard Plus SV Minipreps DNA Purification System kit (Promega)* was used with minor modifications. A single colony was inoculated in 10 ml of LB medium with the respective antibiotics and incubated for 2 days at 28°C in a gyratory shaker at 200 rpm. The culture was centrifuged, using only 2 ml at each time in independent centrifugations for 2 minutes at 5000xg. The remaining protocol was followed according to the manufacturer's instructions.
Plasmid DNA was tested by PCR, to check if *PDS* fragment was inserted in the TRV2 vector, using *PDS* and TRV2 coat protein primers. 20 µl reaction mixtures were prepared with 4 µl of GoTaq Flexi buffer (1X; Promega), 2 µl of $MgCl_2$ (2.5 mM; Promega), 0.4 µl dNTPs mix (0.5mM; Promega), 0.5 µl of each forward and reverse primers (0.25 mM), 0.2 µl of GoTaq G2 Flexi DNA Polymerase (5U/ µl; Promega), 2 µl of plasmid DNA (80 ng/µl) and water to the final volume. PCR was carried out with an initial denaturation step at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 1

min, annealing at 60°C for 1 min and extension at 72°C for 1 min, and one final step of extension at 72°C for 7 min. PCR products were run on an 1% agarose gel and visualized as described in section 2.2.2.

2.2.6. Inoculation experiments

Different experimental conditions were applied for the inoculation of coffee and tobacco plants by testing different protocols of *A. tumefaciens* culture preparation (section 2.2.6.1), methods of inoculation (section 2.2.6.2), OD_{600} values of the bacteria cultures and stages of plant growth.

2.2.6.1 Preparation of *A. tumefaciens* cultures

Three different protocols were tested based on previous studies in *Nicotiana benthamiana*[50,60]and *Solanum lycopersicum*[60], and a protocol from Departamento de Fitossanidade, Universidade Federal de Pelotas, Brazil (Barros D., personal communication).

For the three protocols, single colonies harboring TRV2::*PDS*, empty TRV2 and TRV1 vectors were selected and individually inoculated in 5 ml liquid culture of LB containing 10 mg/ml rifampicin, 50 mg/ml kanamycin and 50 mg/ml gentamicin and grown at 28°C in a gyratory shaker at 200 rpm for about 24 hours. For the remaining part of the procedure, different conditions were applied, as follows:

- Protocol I

1:25 dilution of the primary culture was inoculated into 5 ml of Induction Medium (IM), containing the same antibiotics and 200 µM acetosyringone. The mixture was left to incubate for 24 hours at 28°C in a gyratory shaker at 200 rpm. The induction medium intends to mimic the environment that the pathogen encounters in the host's apoplast, and acetosyringone induces the *vir* genes of *A. tumefaciens* required to transfer T-DNA into the plant.

IM was freshly prepared containing: 50 mM of 2-(4 morpholino)-ethane sulfonic acid (MES), 30 mM of glucose, 18 mM of NaH_2PO_4 and 475 ml of destilled H_2O, pH 5.6. After sterilization, 25 ml of AB salts, containing 374 mM of NH4Cl, 24.4 mM of $MgSO_4.7H_2O$, 40 mM of KCl, 7.2 mM of $CaCl_2.2H_2O$ and 0.39 mM of $FeSO_4.7H_2O$, were added to the IM.

Bacterial cultures were centrifuged twice for 10 minutes at 3000xg. First, the pellet was resuspended in 10 ml, and in half of volume on the second time. Resuspension buffer consisted of 10 mM $MgCl_2$ and 10 mM MES, pH 5.5. OD_{600} was measured in spectrophotometer *UV-1800 (Shimadzu)* and adjusted to the desired value, according to the OD_{600} tested (0.4 and 0.5, Table 2.2). When this value was reached, acetosyringone to a final concentration of 400 µM was added to the TRV1 culture.

- Protocol II

1 ml of the primary culture was subcultured in 20 ml of LB liquid containing the same antibiotics and grown at 28°C in a gyratory shaker at 200 rpm for about 12 hours. The absorbance of each culture was measured as described above. When the OD_{600} tested (0.3, 0.5, 0.6, 0.8, 0.9 and 1, Table 2.2) was reached, bacterial cultures were centrifuged twice for 5 min at 3000xg at room temperature resuspended firstly in 20 ml of *A. tumefaciens* induction buffer, and secondly in 10 ml of infiltration buffer. In between centrifugations, cultures were incubated at room temperature in a gyratory shaker at 50 rpm for 3 hours. After the second resuspension, OD_{600} was measured and re adjusted to the tested value.

All buffers were freshly prepared. The *A. tumefaciens* induction buffer contained 10mM of MES and 200 µl of 200 mM of acetosyringone, pH 5.5. The infiltration buffer had 5 mM of MES, pH 5.5.

- Protocol III

The whole culture was transferred to 50 ml of LB with the same antibiotics and grown at 28°C on a gyratory shaker at 200 rpm until reaching an OD_{600} of 1.0. The culture was then centrifuged for 5 minutes at 10000 rpm and resuspended in Inoculation buffer. The OD_{600} was adjusted to 1.2. The bacterial suspension was incubated for 2 hours at room temperature.

The Inoculation Buffer was freshly prepared containing 10 mM of $MgCl_2$, 10 mM of MES and 200 µM of acetosyringone, pH 5.5.

For the three protocols, the last step was to mix TRV1+TRV2::*PDS* cultures in a 1:1 (v/v) ratio. A vector control containing a 1:1 of TRV1+emptyTRV2 cultures were also prepared. Protocol III was only tested for tobacco.

2.2.6.2 Methods of inoculation and tested conditions

All experiments were carried out with coffee hypocotyls, seedlings and grown plants, and with 3-week-old tobacco plants. Plants without any treatment and plants inoculated with: infiltration buffer or water (mock controls), as well as with TRV1+emptyTRV2 vectors were used as controls. Different methods of inoculation were tested as follows:

- Infiltration (Fig. 2.2. A.I.)

Needleless 2 ml and 5 ml syringes were used to inoculate the abaxial side of the lower leaves with about 0.5 ml of culture, as described by Senthil-Kumar & Mysore (2014)[50]. Three to four leaves per plant were inoculated (Table 2.2.and 2.3)

- Agrodrench (Fig. 2.2. A.II.)

About 6 ml of TRV1+TRV2::PDS culture were dispensed onto the soil near the crown region of the plant, as described by Senthil-Kumar & Mysore (2014)[50].

- Dipping the root (Fig. 2.2. B.)

Roots of plants were briefly washed with water to wipe the soil, and then placed in a recipient containing the bacterial mixture for 15, 30 and 60 min. Plants were then planted in pots with soil.

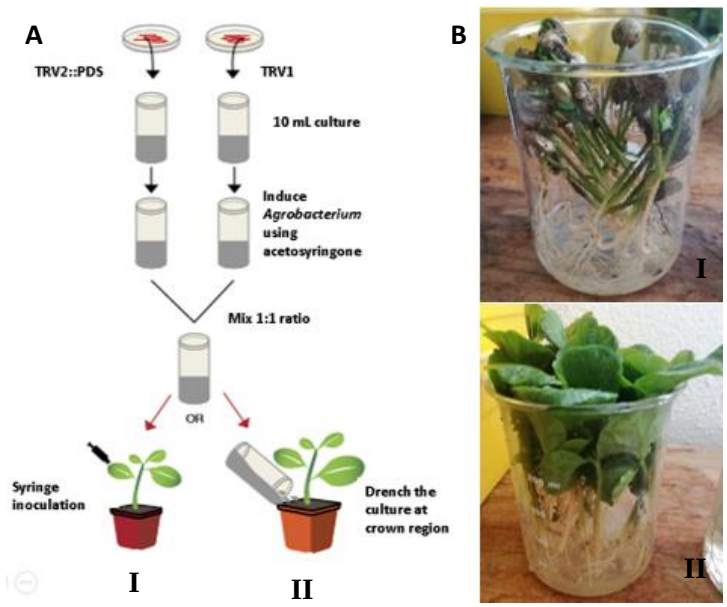


Figure 2. 2 - Methods of inoculation performed for the VIGS experiments. (A) I. Infiltration, II. Agrodrench, adapted from Senthil-Kumar & Mysore (2014)[50]. (B) Dipping of roots of I. hypocotyls, II. seedling.

All tested conditions on coffee and tobacco plants are shown in Table 2.2. and 2.3., respectively.

Table 2. 2 – Inoculation conditions tested on coffee plants. *replicated assay; TL – true leaves; CL – cotyledonary leaves.

Method	Protocol	Developmental Stage		Number of plants				
				OD tested	TRV2: :PDS	empty TRV2	Buffer	W/o treatment
Infiltration	I	Grown		0.4	6	2	2	2
				0.3	10	2	2	2
	II	Seedlings	TL	0.8	12		12	12
				0.6	12			
				0.6	10		7	
			CL	0.8	12		12	12
				0.6	12			
				0.6	10		7	
Agrodrench	II	Grown		1	10	2	1	2
		Seedlings	CL	1	15*		13	12
				1	23		10	
		Hypocotyls		1	20*		14	12
				1	13		10	
Dipping	15 min	Seedlings	CL	0.9	10			12
		Hypocotyls		0.9	10			
	30 min	Seedlings	CL	0.9	10		10	
		Hypocotyls		0.9	10		10	
	60 min	Seedlings	CL	0.9	10			
		Hypocotyls		0.9	10			

Table 2. 3 – Inoculation conditions tested on tobacco plants.

Method	Protocol	Number of plants				
		OD tested	TRV2: :PDS	empty TRV2	Buffer	W/o treatment
Infiltration	I	0.4	2	2	2	1
		0.5	4			1
	II	0.6	9		2	1
		0.5	10	2	1	1
	III	1.2	12		2	2
Agrodrench	I	0.5	2		1	1
	II	1	10	2		1

Two types of conditions for maintaining the plants after inoculation were tested: half were kept in greenhouse conditions (temperatures measured from May until August), and the other in a phytotron FitoClima S600 (Aralab) at 22°C with 14h light/10h dark.

2.2.6.3 Detection of TRV2 coat protein fragment on inoculated plants

In the absence of photobleaching symptoms, the presence of the viral vector within inoculated plant cells was assessed. Samples were collected from a total of 4 coffee plants used in VIGS experiments: 1 leaf pair from a non-inoculated plant, and 3 new emerging leaf pairs from inoculated plants. For

tobacco, samples were collected from 6 plants: 1 leaf pair from a non-inoculated plant, 2 inoculated leaf pairs and 3 new emerging leaf pairs from inoculated plants. All collected leaves were used for RNA extraction and cDNA synthesis as explained in section 2.1.3. PCR using the TRV2 primers was performed to detect if the vector was efficiently transferred to the plant during the inoculation process. PCR conditions were: 20 µl reaction mixtures with 4 µl of GoTaq Flexi buffer (1X; Promega), 2 µl of $MgCl_2$ (2.5 mM; Promega), 0.2 µl dNTPs mix (0.04mM; Promega), 0.5 µl of each forward and reverse primers (0.25 mM), 0.2 µl of GoTaq G2 Flexi DNA Polymerase (5U/ µl; Promega), 1µl of cDNA (50 ng/µl) and water to the final volume. PCR was carried out with an initial denaturation step at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 60°C for 1 min and extension at 72°C for 45 s, and one final step of extension at 72°C for 7 min. PCR products were run on an agarose gel and visualized as described in section 2.2.2.The obtained fragments were purified using the *Wizard SV Gel and PCR Clean-Up System kit (Promega)*, according to the manufacturer's instructions and sent to sequencing to STAB VIDA.

3 RESULTS

3.1 Gene expression analysis

3.1.1 Cytological analysis of rust infection

A cytological analysis was performed along the key stages of the infection process to monitor the success of the infection and differences in fungal development between both coffee-rust interactions, thus to validate the time-points selected for gene expression analysis and better correlate results. At 24 hpi, the percentages of germinated uredospores and appressoria formed (Fig. 3.1.) on stomata were similar in susceptible and resistant coffee leaves (Table 3.1.).

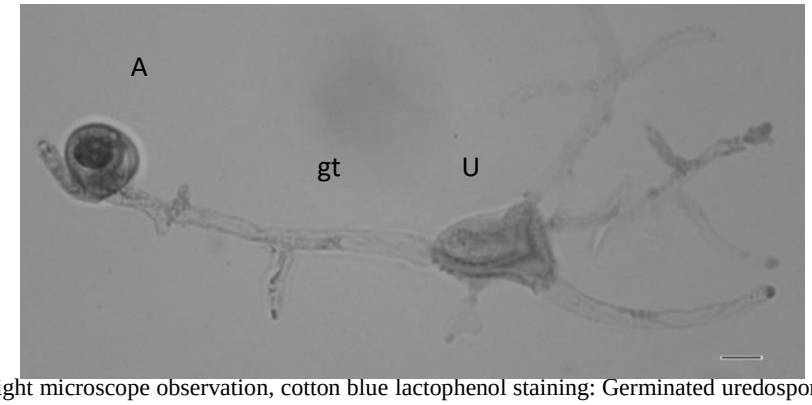


Figure 3. 1 – Light microscope observation, cotton blue lactophenol staining: Germinated uredospore (U) with germ tube (gt) and appressorium formed (A) differentiated on the lower surface of the leaf (compatible interaction), 24 hpi. (bar=10µm)

Table 3. 1 – Percentage of germinated uredospores and appressoria formed by *H. vastatrix* (isolates 71 and 1427) in Caturra, at 24 hpi. For statistical analysis of cytological data, Student's t test was applied and non-significant statistic differences were obtained between the two isolates in percentage of both germinated uredospores and appressorium.

H. vastatrix structures	Isolate 71	Isolate 1427
% germinated uredospores	61 ±9	53 ± 8
% appressorium	49 ± 7	47 ± 8

Microscopic examination of fungal growth inside the host tissues were monitored during the time course of the infection. At 24 hpi, uredospores of both races had germinated and developed an appressorium over stomata, with no difference between the two. Fungal infection of leaf tissues significantly differed between incompatible and compatible interactions at 72 hpi ($P \leq 0.001$) and 96 hpi ($P \leq 0.01$), as shown by quantification of fungal differentiated structures (Fig. 3.2.). In the incompatible interaction, fungal development was slower from 48 hpi onwards and at 96 hpi had a few haustorial mother cell (HMC) with haustorium structures (2%), being the appressorium the predominant stage (Fig. 3.3.). In contrast, fungal development of the virulent *H. vastatrix* strain reached anchor and HMC with haustorium stages at 48 hpi.

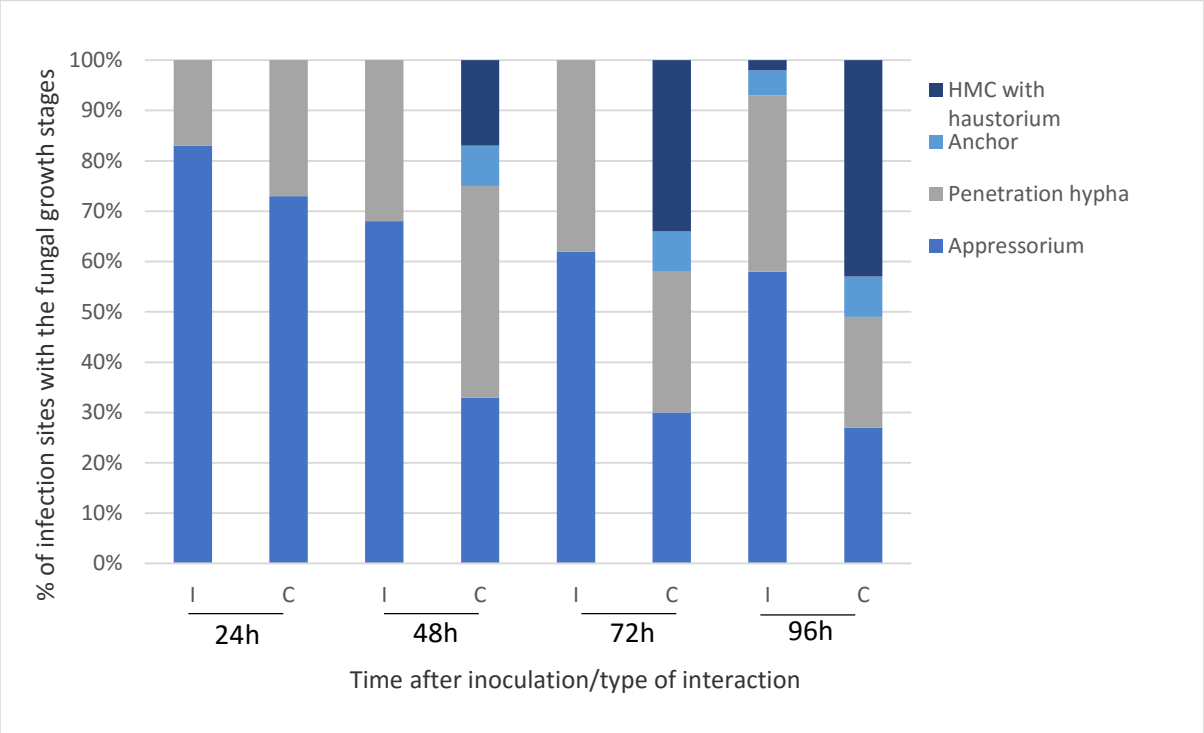


Figure 3. 2 – Percentage of infection sites and respective infection structures differentiated by *H. vastatrix* at different times after inoculation in the incompatible and compatible interaction (I and C, respectively). HMC - haustorial mother cell

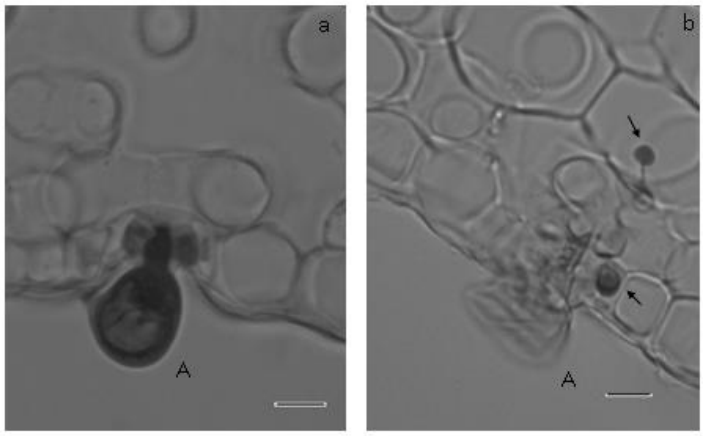


Figure 3. 3 – Light microscope observations, cotton blue lactophenol staining: colonization of leaf tissues by *H. vastatrix* in the (a) incompatible and (b) compatible interactions, at 96 hpi. (a) Infection site showing an appressorium (A) over the stomata; (b) Infection site showing an appressorium (A) over stomata and HMCs with

haustoria (arrow) within a subsidiary stomatal cell and a cell of the second layer of spongy parenchyma. (bar= 10µm).

3.1.2 Expression profiles of candidate genes for coffee resistance

Compatible and incompatible *C. arabica – H. vastatrix* interactions were studied by the expression analysis of eleven genes related with resistance to assess differential patterns between the responses of resistance and susceptibility in coffee. However, only 7 genes could be analyzed (Fig. 3.4.and Supplementary Fig.3). qPCR analysis of the remaining genes (*PR1b, LRR1241, LRR830.5, β-Glu 19724*) showed multiple peaks in the melting curve. Although different conditions were tested, including re-design of primers, amplification of these genes could not be optimized and thus had to be abandoned.

For the analyzed genes, no statistical differences between incompatible and compatible interactions were obtained during the infection time course. However, some differential patterns of expression and differences in expression levels were noted.

The *GL22853* and *Asp23673* genes showed to be down-regulated at all time points of the infection process, except at 96 hpi for the *Asp23673* gene. In most time points no significant visual difference was observed between both interactions. However, a higher down regulation in expression was detected in compatible interaction at 24 hpi for the *Asp23673* gene (compatible: -14.9 ± 0.029, incompatible: -5,49) and also at 72 hpi in the incompatible interaction (compatible: -5.73 ± 0.22, incompatible: -11.53 ± 0.095). Interestingly, for *GL22853* a higher level of repression was also observed at 72 hpi in the resistance response compared to the incompatible (compatible: -2,24 ± 0.4, incompatible: -3,62 ± 0.33). Remarkably, there is statistical difference in the expression between 24 hpi and 48 hpi in the incompatible reaction and 24 hpi and 96 hpi in both interactions in the *GL22853* gene, and also between 24 hpi and 48 hpi in the compatible interaction for the *Asp23673* gene.

On the other hand, *GL18058* and *Asp203* genes had no significant expression throughout the infection process, except for an upregulation in *Asp203* at 72 hpi in the susceptibility response (12,49 ± 4). Although not statistically significant, it was also possible to detect some differential pattern of expression between resistant and susceptible samples at 24 hpi, and between 24 and 48 hpi in the resistant response, in particular in the *GL18058* gene. (compatible: 1,13 ± 0,46 [24 hpi], incompatible: -1,01 ± 0,22 [24hpi], 1,55 ± 0,43 [48 hpi]).

Remarkably, the recognition and pathogenesis related genes (*RLK, PR1* and *PR10*) were up-regulated at all time points of the infection process in both interactions, except for 96 hpi in *RLK*. When comparing both responses, a significantly higher increase in expression was obtained at 24 hpi for the *PR1* gene and at 48 hpi for *PR10* and *RLK* in the resistant response (incompatible – *PR1*: 49,04 ± 42, *PR10*: 84,0 ± 99.8, *RLK*: 16,12 ± 4.7; compatible - *PR1*: 4,60 ± 2.8, *PR10*: 23,97 ± 32.7, *RLK*: 2,35 ± 3). In the compatible reaction, expression of all these genes reached a maximum at 72hpi (*PR1*: 14,02 ± 16.8; *PR10*: 71,00 ± 75.3; *RLK*: 22,11 ± 3.29), which was significantly higher when compared with the incompatible one (*PR1*: 1,62 ± 0.75; *PR10*:19,87: ± 12.8; *RLK*: 2,76 ± 2.35).

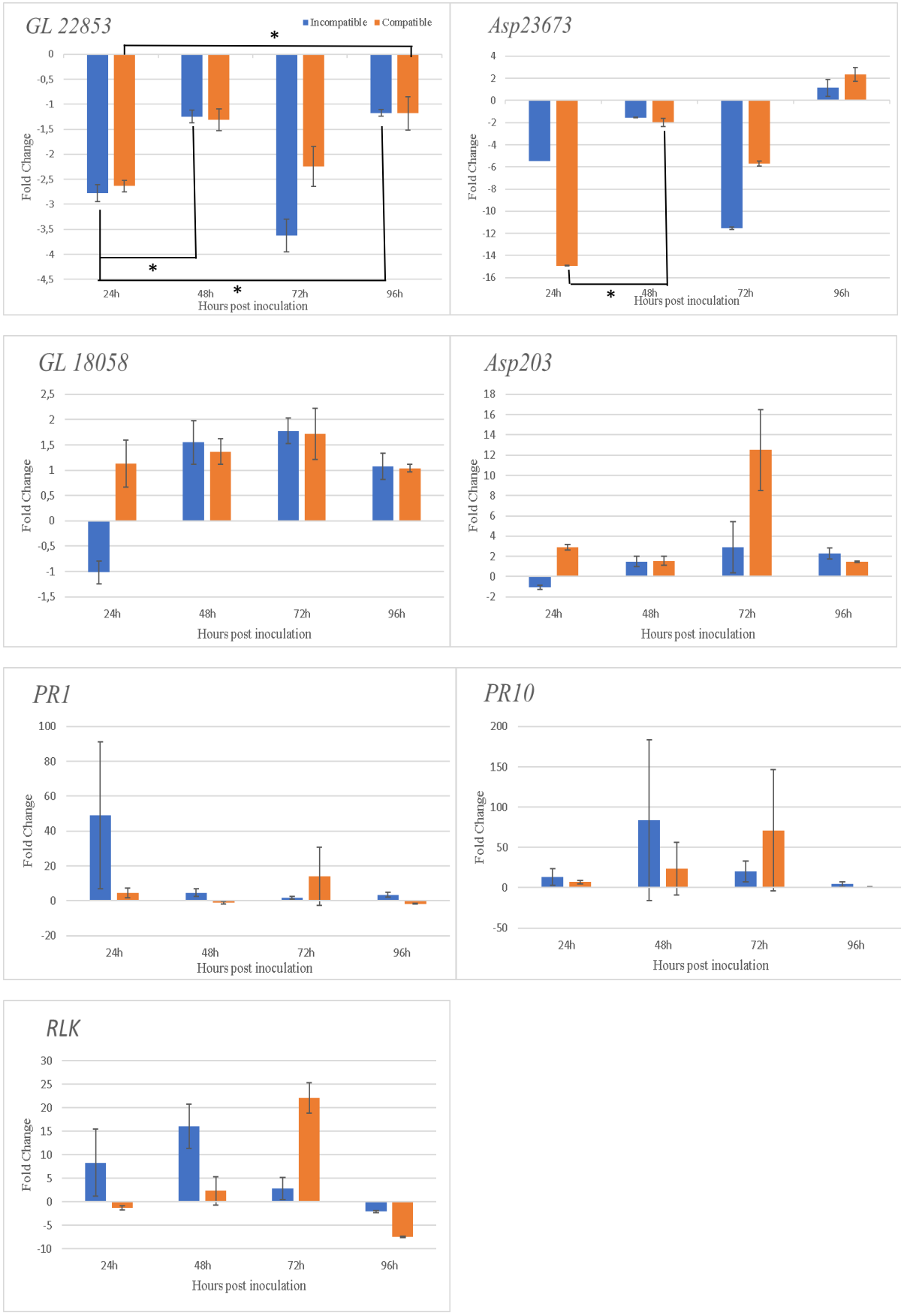


Figure 3. 4 – qPCR expression analysis of studied genes. Relative expression patterns of: *GL22853*, *Asp23673*, *GL18058*, *Asp203*, *PR1*, *PR10* and *RLK*. * represents statistical difference at P=0.05.

3.2 Virus-induced gene silencing

3.2.1 Selection of transformants

In general, transformation using a 3:1 molar ratio of insert:vector ligation reaction produced better results than the 1:1 ratio. For *E. coli* transformed with *pGEM-T Easy* vector with the *CaPDS* insert, 8 white colonies were selected and tested by colony PCR. From these, 4 colonies were confirmed to be positive (50%). For transformants with *pGEM-T Easy* vector with the *NrPDS* insert, 8 white colonies were selected, and 1 was PCR positive (12.5%). The expected bands were visualized on the agarose gel (Fig. 3.5.) and confirmed by sequencing (Supplementary Fig. 4). The selected transformants used in the preparation of TRV2::*PDS* vectors are highlighted in Figure 3.5.

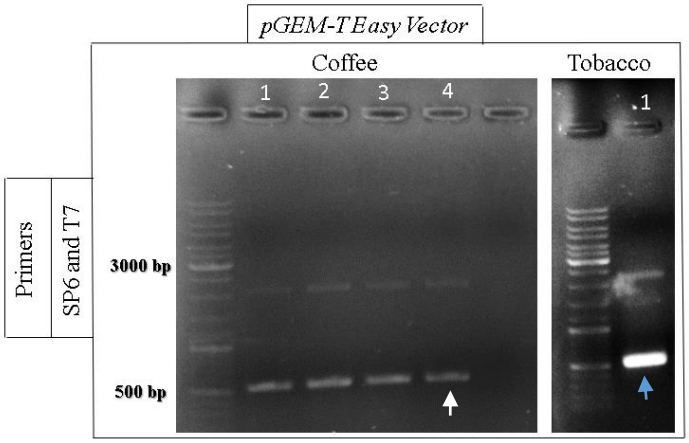


Figure 3. 5 – *pGEM-T Easy* vector constructs with the *PDS* insert from coffee (*CaPDS*) and tobacco (*NrPDS*): presence of the expected band on electrophoresis gel. White and blue arrows highlight the transformants with the insert from *CaPDS* and *NrPDS*, respectively, used for the remaining work. Ladder: 1 kb.

For *A. tumefaciens* transformed with TRV2::*CaPDS*, 24 colonies were selected and tested by colony PCR. From these, only 3 showed to be positive colonies (12.5%). For transformants with the TRV2::*NrPDS* vector, 16 colonies were tested and only 2 were positive (12.5%).
From these positive colonies, plasmid DNA was extracted and a confirmation PCR for the presence of *PDS* was performed. Detection of *PDS* fragments on the electrophoresis gel for those colonies/plasmid DNA is indicated with a white arrow on Figure 3.6. In the plasmid DNA of colony 3 from coffee and of colony 2 from tobacco, it was not possible to confirm the *PDS* presence.

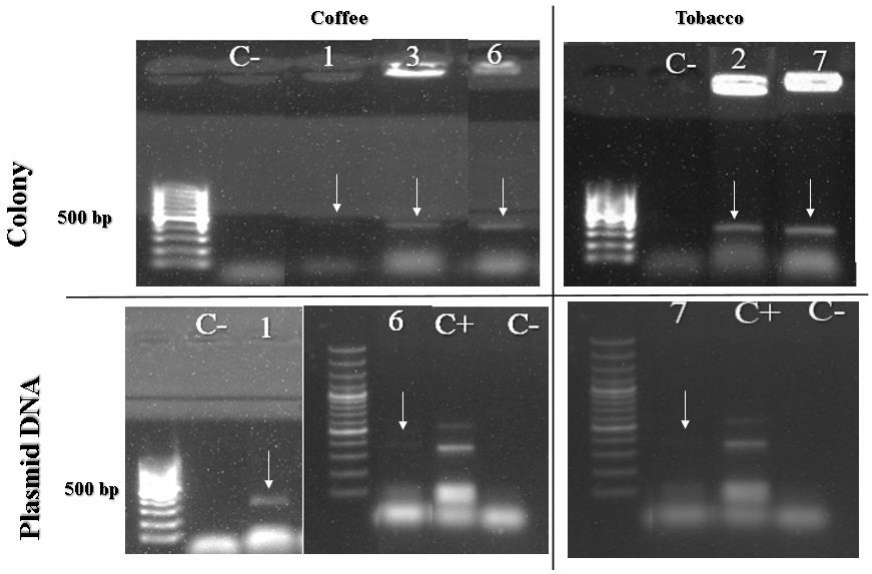


Figure 3. 6– Presence of the expected band (*PDS*, 345 bp) on electrophoresis gel, for colony and plasmid DNA of constructs with the *PDS* insert from coffee or tobacco. Expected bands are indicated with white arrows. C- represents the negative control of the PCR amplification, without template. C+ represents the positive *PDS* control, using cDNA from coffee. Ladder: 100 bp plus.

3.2.2 Detection of viral infection

To assess if the viral plasmid was introduced into the plant cells by *A. tumefaciens* infiltration, amplification of the TRV2 fragment on the infiltrated plants was tested.

From the 4 coffee plants selected, only one of the samples, corresponding to new emerging leaves, showed the expected band corresponding to TRV2 fragment (Fig. 3.7.). From the 6 tobacco plants selected, also one of the samples, corresponding to the infiltrated leaf pair, had the expected band (Fig. 3.7.). This band corresponds to a 401 bp TRV2 fragment, as shown in the positive (+) control (from TRV2 plasmid) and proved by sequencing.

The upper band (~700 bp) from both tobacco and coffee, represented by a blue arrow in Figure 3.7, was sent for sequencing. It was shown to correspond by 99% homology to a partial sequence of the chloroplast genome of *N. undulata* and *C. arabica*, respectively (Supplementary Fig. 5).

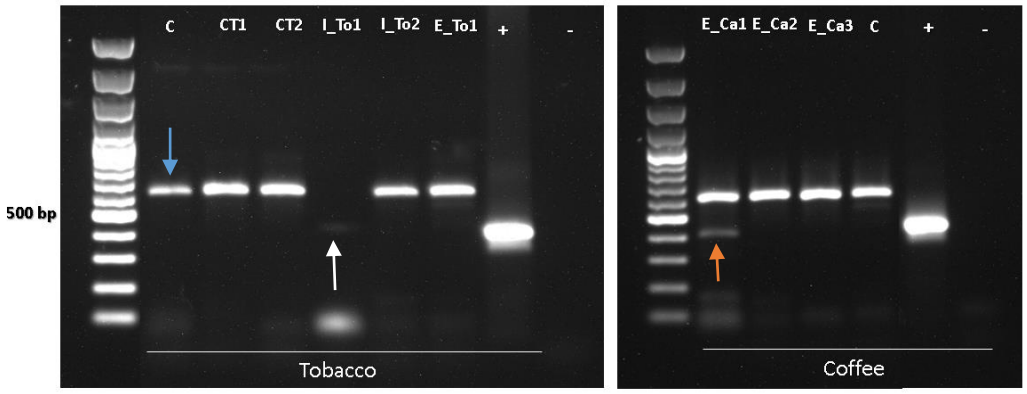


Figure 3. 7– Presence of the expected band (TRV2) on electrophoresis gel after amplification of the fragment from VIGS experiments of tobacco and coffee plants, to assess if the viral plasmid was introduced into the plant cells by *A. tumefaciens* inoculation. White and orange arrows correspond to the expected band, present in samples of infiltrated leaves in tobacco (I_To1) and new emerging leaves of coffee (E_Ca1), respectively. C corresponds to control samples of leaves that were not

inoculated, CT1-2 corresponds to control samples of leaves inoculated with TRV1+emptyTRV2, I_To1-2 corresponds to samples of leaves inoculated with TRV1+TRV2::PDS, E_To1and E_Ca1-3 corresponds to samples of emerging leaves of plants inoculated with TRV1+TRV2::PDS, in tobacco and coffee plants respectively. Negative control of the PCR, without samples (-). Positive control, using plasmid DNA from TRV2 empty vector (+). Ladder: 100 bp plus.

3.2.3 Observation of infiltrated plant leaves

The expected *PDS* suppression phenotype was not visible after agroinoculation of both *N. rustica* and *C. arabica* plants. Careful observation of putative effects of *PDS* silencing showed no difference between inoculated and control plants. This demonstrates that inoculation of TRV2::C*aPDS* and TRV2::*NrPDS* was not successful in achieving silencing of *PDS* in the tested plants. The expected symptoms from silencing would be photobleaching, putatively caused by the absence of the photoprotective carotenoid pigments that require phytoene desaturase for their synthesis. For the infiltration assay on grown coffee plants, no difference was seen between infiltrated and control plants throughout time (Fig.3.8. A and C). Also, no signs of silencing of PDS were noted (Fig. 3.8. A and B).

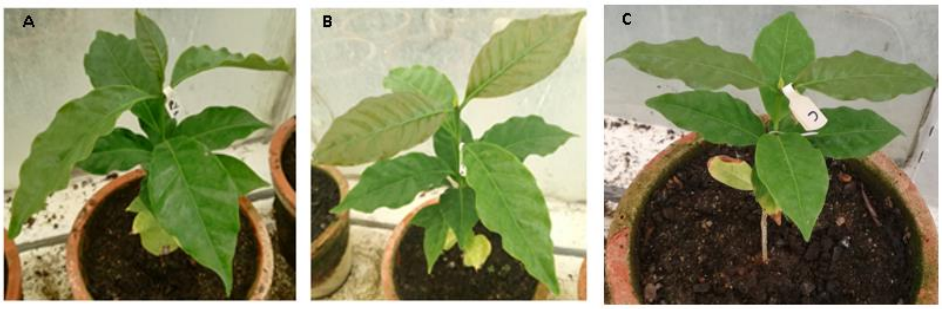


Figure 3. 8– Same plant (A) on the day it was infiltrated, (B) 3 weeks after infiltration and (C) a control plant that was not subjected to any treatment.

The assays performed on coffee plants at different stages of development also did not show the photobleaching phenotype. In terms of development, plants that were at the stage of first leaves and seedling grew normally, as seen in comparison with their respective control plants (without treatment). However, when plants were at the hypocotyl stage only about 40% developed into seedlings, independently of the method of infiltration applied, whether by dipping or agrodrench, and also independently of the inoculum, TRV2::*PDS* or buffer (Fig. 3.9. A and B).

Figure 3. 9- Development of coffee plants at the hypocotyl stage (A) on the day of the assay and (B) 2 weeks after agrodrench.

For tobacco plants, also no signs of photobleaching phenotype were detected. Plants that were infiltrated using a 5 ml syringe suffered lesions on the leaves, ending on their progressive necrosis throughout time. When using a 2 ml syringe, most plants recovered from eventual lesions caused by

infiltration (Fig. 3.10. A). However, tobacco plants that were infiltrated with TRV2, whether the vector was empty or with the insert, lesions tended to become bigger throughout time. Control plants infiltrated with buffer or that were not subjected to any treatment had a green natural color and no signs of lesions (Fig. 3.10. B).

Figure 3. 10- Effects of infiltration in tobacco plants. (A) Leaf with lesion one month after infiltration with TRV2::*PDS*. (B) Lesions on leaves of plants infiltrated with TRV2::*PDS*, represented with a green arrow, and empty TRV2, represented with an orange arrow. No damage is seen in control plants: Without any treatment (blue arrow) and infiltrated with buffer (yellow arrow).

On both coffee and tobacco plants, the outcome of the different assays followed the described patterns, independently of the protocol used.

4 DISCUSSION

4.1 Expression profiling of resistance candidate genes in coffee-rust compatible and incompatible interactions

Cytological analyses were performed to monitor the fungus' development during the infection process between both coffee-rust interactions and to validate the time-points selected for qPCR analysis for gene expression study. In this work, the cytological observations of coffee leaves revealed that in the incompatible interaction the fungus slowed its growth when it reached the appressorium stage, barely developing the subsequent infection structures, possibly due to the host hypersensitive response (HR) as previously described[24,38,39]. On the contrary, in the compatible interaction the fungus kept its growth without apparent inhibition. So, while at 24 hpi fungal development is very similar in both interactions, from 48 hpi onwards striking differences in fungal infection begin to be revealed between both interactions, showing statistically significant variation at 72 and 96 hpi.

In this study, the expression profiles of 11 genes were analyzed by qPCR. Different conditions for qPCR amplification of these genes were tested, however it was not possible to optimize it for 4 of them, making it impossible to study their role in this interaction. The 7 genes studied were *GL18058*, *GL22853*, *Asp23673*, *Asp203*, *PR1*, *PR10* and *RLK*.

Genes *GL18058* and *GL22853* encode proteins of the germin family. Germin-like proteins (GLP) belong to a superfamily of proteins called 'cupin'. These proteins are involved in various enzymatic activities, including superoxide dismutase[36], also acting as receptors in signal transduction. Previous research using transgenic plants and gene expression evaluation suggests that GLP are related to the increase of resistance to various types of pathogens in different plant species, due to enhancement of oxalic acid metabolism, raise of endogenous ROS, reinforcement of the cell wall during pathogen

attack and accumulation of H_2O_2[61-63]. Silencing of a GLP gene in Asian soybean led to the decrease of resistance to rust, with an increase in the percentage of haustoria when compared with control plants, while overexpression led to an induction of resistance with a reduction of rust symptoms per leaf area[36]. These proteins have been identified during the resistant response of coffee to *H. vastatrix* [34,35], showing a higher expression level in the resistant than in the susceptible samples. However, in the present study no significant expression was observed in *GL18058* and the *GL22853* gene was repressed in the time points studied of the infection process. Yet, in *GL18058* gene a contrary profile was detected between incompatible and compatible interactions at 24 hpi, and an alteration of the expression profile was also detected in the resistant samples, from time-point 24 to 48 hpi, suggesting some kind of activation during the incompatible interaction, although showing low levels of expression. Therefore, it would be interesting to study intermediate time points to better understand this pattern. On the other hand, down-regulation of *GL22853* gene was found statistically significant between different time points for both interactions, which might indicate some role in plant reaction but not directly related with resistance.

Genes *Asp23673* and *Asp203* code for aspartic proteases (APs). APs have been studied in many different plant species, but their biological functions are not well characterized. These proteins were described as involved in stress responses in tobacco and tomato leaves by degrading PR proteins which had accumulated in response of plants during stress situations, thereby preventing their overaccumulation[64]. Nepenthesin is an acid proteinase secreted by *Nepenthes* species that belongs to APs. These proteins have been associated to protein catabolism, sorting, degradation and folding processes, and have been identified by whole genome sequencing as contributors to sheath blight disease resistance in rice[65]. Nepenthesin proteins have been identified as important for coffee's resistance to *H. vastatrix*[35], showing a higher expression in the resistant reaction when compared to susceptible samples[34]. In the present study, the A*sp203* gene is weakly induced in all time points, except at 72 hpi in the compatible interaction, where the expression is higher. A previous study in *Arabidopsis thaliana* by Yao et al. (2012)[66] demonstrated that *Aspartic protease in guard cell 1 (ASPG1)* gene expression is induced by abscisic acid (ABA) under drought conditions, not having a known involvement in responses to biotic stresses. Nevertheless, it would be interesting to study the expression of *Asp203* gene in the first hours of the infection process, when the fungus is penetrating the stoma, since this gene is an *ASPG-like* and may play some defense role in the guard cell. The A*sp23673* gene is repressed, with a significant difference between 24 and 48 hpi in the compatible interaction, except at 96 hpi, which is weakly activated.

The results obtained in this work for these four genes do not corroborate previous ones. Since GL and AP proteins previously studied showed to be overexpressed during the response against rust[35], it is possible that these genes might have been differentially expressed in an earlier time of the infection process, and not expressed in the time points studied in this work. Another explanation, is that these genes might be negatively regulated to inhibit fungal infection/growth. Also, since either the GL and AP genes studied belong to large families with different biological functions, it is possible that these specific genes do not play an import role in host-pathogen interaction. Nevertheless, and taking into account the APs role described by Simões & Faro (2014)[64] the expression profile of *Asp23673* gene could somehow be related with the expression pattern found for the pathogenesis-related (PR) genes studied.

In this work, two PR genes (*PR1*and *PR10*) were studied. *PR1* are proteins known to be ubiquitous across plant species, which transcription is activated upon accumulation of salicylic acid (SA), due to pathogen attack. Overexpression of *PR1* in transgenic tobacco plants was shown to lead to an increase

of pathogen resistance, suggesting that these proteins are antimicrobial[67]. In the susceptible and resistant varieties of coffee to *C. kahawae*, the expression patterns of *PR1* gene were very similar for both varieties, being up-regulated from 12 hpi onwards, but in a greater magnitude in the resistant variety[38]. This accumulation on both varieties corresponds to basal defense responses regulated by SA. Previous studies of *Coffea* spp- *H. vastatrix* showed that *PR1* gene is activated around 24 hpi in the incompatible interaction[34]. Corroborant to those results, in this work, *PR1* gene was highly activated at 24 hpi in the incompatible interaction, decreasing in expression immediately afterwards, which suggests that an early regulation of this gene expression is involved in halting fungal development. On the contrary, in the compatible interaction, an activation peak of smaller magnitude is observed, but only at 72 hpi, showing a later and weaker reaction of the plant to infection.

On the other hand, accumulation of jasmonic acid (JA) during pathogen attack has been described, controlling gene transcription of other genes, such as *PR10*[38]. Although *PR10* functions are poorly understood, it is known that some PR10 protein members have antimicrobial activity, DNase and/or RNase activity, while others control the phenylpropanoid and flavonoid biosynthesis and their transport to, for example, the cell wall[38,68,69]. In the interaction of coffee with *C. kahawae*, *PR10* gene showed differences of expression levels in both coffee varieties, in time and magnitude. In the susceptible reaction, this gene showed a maximum value at 72 hpi corresponding to the necrotrophic fungal growth, while in the resistant reaction, high expression level of this gene was seen at the beginning of fungal penetration, at 48 hpi[39]. Although the interaction studied in this work is different, comprising a fungal biotrophic development, the expression of *PR10* gene has a very similar pattern. The increase in *PR10* expression was seen at 48 hpi in the incompatible interaction, although fungal penetration in this case begins before 24 hpi, and at 72 hpi in the compatible interaction. While in the susceptibility reaction, once again, *PR* genes are also activated but only later, probably too late to stop pathogen development, in the resistant reaction *PR10* activation seems related with resistance, but probably not directly to the earliest attempt to halt pathogen development. Interestingly, the time at which accumulation of *PR10* and *PR1* transcripts seem to stop and activation of *Asp23673* seems to start are coincident, which could reflect a role of this AP as described by Simões and Faro (2014)[64] in degrading overaccumulated *PR*s.

RLK are transmembrane kinases that perceive diverse signals and stimuli from the environment, participating in processes such as hormone perception, regulation of development, and some may be involved in resistance signaling pathways[37,70]. In a previous study using another incompatible interaction of coffee-*H. vastatrix*, *RLK* gene presented two activation peaks: a first one corresponding to the time of appressoria differentiation and penetration hypha formation (6-12 hpi), and a second one when anchors and haustoria mother cells differentiated (24 hpi)[37]. In our work, overexpression was observed since 24 hpi, with a peak of activation at 48 hpi in the incompatible interaction, corresponding only to appressoria differentiation and penetration hypha formation in this interaction. This could be explained by a slower interaction development in our study, but since the time-courses studied are different a direct comparison cannot be made. Also, gene expression patterns obtained in our work nearly match the ones obtained in studies of interactions of coffee hypocotyls with *C. kahawae*, for the compatible interaction[39]. In our study, *RLK* expression in the compatible interaction was repressed at 24 hpi, but at 72 hpi reached a maximum expression, with even higher magnitude than in the incompatible interaction. These results confirm the latter response of the susceptible host to the pathogen attack. Moreover, *RLK* gene up-regulation does not seem to be required in later stages of infection, since repression was observed in both interactions at 96 hpi.

Overall, these results confirmed that the signaling and defense related genes studied seem to be involved in coffee defense responses, and that are up-regulated in both compatible and incompatible interactions, but at different timings of the infection process. In the resistance response, these genes are early activated in the time course of the infection and the plant is able to arrest the fungus growth at the appressoria phase. In contrast, in the compatible interaction, these genes showed differential expression only at 72 hpi, possibly too late to stop the fungus growth, therefore enabling the infection process to be completed. In both interactions, 96 hpi seems to be a time point in which all studied genes have no significant relevance in the infection process.

Based on the results obtained, *Asp23673, PR1, PR10* and *RLK* genes would be most interesting to be further studied, for instance along a more comprehensive infection time-course, and within more detailed functional approaches, such as VIGS, to better understand their role in coffee resistance.

4.2 Development of Virus induced gene silencing (VIGS) for coffee

There is the need to develop efficient gene functional approaches to understand various biological processes in higher plants. TRV-based VIGS is a quick tool that silences or down-regulates the expression of any target gene, in which the infected plants employ the innate PTGS as a defense mechanism against the TRV. In this work, the silencing of *PDS* through VIGS was inefficient, since plant inoculation with TRV2::*PDS* did not show the expected white chlorosis in the new emerging leaves of both coffee and tobacco plants. In previous studies, inoculated *N. benthamiana* plants had the *PDS* endogenous transcript levels reduced by 95%[71].

There are a wide range of parameters that may influence the practical application of this system. A major problem faced in this work was the delivery of TRV1 and TRV2 constructs through syringe inoculation into plants of both species. This method created large lesions in some of the inoculated leaves when 5 ml syringe were used, being preferable to use 2 ml syringes. Lesions were also detected in most tobacco plants infiltrated with TRV2 vector, either empty or with the insert, probably because TRV can cause mild virus-associated phenotypic changes[50]. The other two used methods did not cause any visible damage to the plants, however it also did not result in the photobleaching phenotype.

Moreover, plant-virus interaction depends on environmental conditions, varying between species[55], and *Agrobacterium*-mediated transient gene expression and virus multiplication are also known to be influenced by temperature[72]. In this experiment most plants were kept at greenhouse conditions, in which temperature ranged from 23°C-32°C, depending on the weather. The remaining plants were kept in a phytotron, with controlled conditions: 22°C, 14h light/10h dark. Other and more stable temperature conditions could have been tested. On the other hand, the conditions that are more suitable for plant growth may not be the best for *A. tumefaciens*-mediated delivery for the virus vectors, since coffee requires tropical conditions,

Also, in this work, 3 different protocols were tested, in which the main differences were the buffers used to inoculate *A. tumefaciens* and different tested OD_{600} values (between 0.3 and 1.2) of bacterial suspensions prior inoculation. Besides, different stages of development of coffee plants were subjected to the inoculation process. These tested conditions did not seem to influence the success of the VIGS procedure and subsequently the appearance of the photobleaching phenotype. However, within the plant developmental stages tested, the hypocotyl stage does not seem to be a good option to inoculate, since independently of the inoculation method tested in this work, its growth seemed to be affected, probably due to manipulation.

Another difficulty of using VIGS is that the vectors can be limited by species specificity since the constructs are based on natural plant viruses, being TRV vector the most used VIGS vector so far and has a broad host range[73], being commonly used in *N. benthamiana*[50]. However, silencing by TRV-VIGS has never been previously tested in the species used in this work. Another vector could be used for coffee plants, or a specific vector should be designed using the backbone of Coffee ringspot virus (CoRSV), which is a single-stranded RNA plant-adapted virus that causes great concerns in coffee cultures[74]. This would be possible if nonstructural proteins from the virus were replaced with multiple cloning sites to insert fragments of the gene to silence[50].

In addition, although not likely, the *PDS* fragment could have been lost after ligation[50]. However, it was possible to confirm through PCR that the constructs had the *PDS* fragment inserted, although some PCR amplification reactions were not reproductible. It is also possible for the inoculated plants to have a downregulation of the *PDS* gene which is not enough to be visible as the expected phenotype. In this work, no differences in the expression levels of the *PDS* gene between inoculated and non-treated plants were noted in qPCR analyses (data not shown).

Also, confirmation of the viral vector presence in inoculated plants, through PCR amplification of the TRV2 coat protein fragment was obtained, while in control plants there was no amplification. In this PCR reaction, an unexpected ~700 bp fragment was also amplified, for both tobacco and coffee samples. Upon sequencing, this fragment was identified as a partial gene encoding a photosystem II protein D1 in the chloroplast genome, which shows homology in the primer annealing sequence with the TRV2 coat protein. An explanation for this amplification is the presence of retrotransposons in the chloroplast genome, proven to exist by the detection of insertions of chloroplast DNA in the nuclear genome[75,76], and resulting from past infections by viruses[77]. Nonetheless, these results show that the virus entered in the cells of some inoculated plants, but it was not efficient, which could have been caused by a problem in the co-delivery of the TRV1 vector, essential for viral replicon and movement within the host plant, or disruption or loss of the *PDS* insert during the process. Further studies should be conducted and other approaches tested, since developing a successful VIGS system in coffee plants would be a major tool to explore the genes putatively involved in resistance to rust.

In conclusion, qPCR data have indicated that some of the studied genes are differentially expressed in response to the virulent or avirulent rust pathogen infection. Future work will aim at better studying those genes and finding additional genes that putatively have a role in coffee-rust interaction, and develop a successful VIGS system capable of silencing genes in coffee plants. Integration of the results of such studies would allow silencing of those genes to understand their role at the time point of the infection process in which they are involved. Altogether, improvement of the selection for resistance to CLR would be possible, since it would allow allele-mining and the development of molecular marker[24,35].

5 REFERENCES

1. Cronquist, A. (1988) *The evolution and classification of flowering plants*. New York Botanical Garden, New York.

2. Bridson D. (1994) Additional notes on *Coffea* (Rubiacea) from tropical East Africa. *Kew Bull* **49**: 331–342.

3. Anthony F., Diniz LEC., Combes MC., Lashermes P. (2010) Adaptive radiation in *Coffea* subgenus *Coffea* L. (Rubiaceae) in Africa and Madagascar. *Plant Systematics and Evolution* **285**:51–64. doi:10.1007/s00606-009-0255-8.

4. Davis AP. (2003) A new combination in *Psilanthus* (Rubiaceae) for Australia, and nomenclatural notes on *Paracoffea*. *Novon* **13**: 182–184. doi:10.2307/3393514

5. Davis AP., Govaerts R., Bridson DM., Stoffelen P. (2006) An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot. J. Linn. Soc* **152**: 465-512.

6. Davis AP., Tosh J., Ruch N., Fay M.F. (2011) Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot. J. Linn. Soc* **167**: 357–377.

7. Talhinhas P., Batista D., Diniz I., Vieira A., Silva DN., Loureiro A., Tavares S., Pereira AP., Azinheira HG., Guerra-Guimarães L., Várzea V., Silva MDC. (2017) The coffee leaf rust pathogen *Hemileia vastatrix*: one and a half centuries around the tropics. *Mol. Plant Pathol.* **18**:1039–1051. doi:10.1111/mpp.12512.

8. Mishra MK., Slater A. (2012) Recent Advances in the Genetic Transformation of Coffee. *Biotechnol. Res. Int.* **2012**:1-17. doi:10.1155/2012/580857.

9. Rodrigues CJ., Bettencourt AJ., Rijo L. (2013) Races of the Pathogen and Resistance to Coffee Rust. *Annu. Rev. Phytopathol.* **13**(1):49-70. doi:10.1146/annurev.py.13.090175.000405.

10. Clifford MN., Willson KC. (1985) Botany, Biochemistry and Production of Beans and Beverage. *AVI Publishing Company, Inc*. Westport, Connecticut.

11. DaMatta FM. (2004) Ecophysiological constraints on the production of shaded and unshaded coffee: A review. *F. Crop. Res.* **86**(2-3):99-114. doi:10.1016/j.fcr.2003.09.001.

12. Labouisse JP., Bellachew B., Kotecha S., Bertrand B. (2008) Current status of coffee (*Coffea arabica L.*) genetic resources in Ethiopia: Implications for conservation. *Genet. Resour. Crop. Evol.* **55**(7):1079-1093. doi:10.1007/s10722-008-9361-7

13. Wrigley G. (1988). *Coffee. Tropical Agriculture Series*. Longman Scientific & Technical, New York.

14. Waller JM., Bigger M., Hillocks RJ. (2007) *Coffee Pests, Diseases and their Management*. CABI Publishing, Wallingford, UK.

15. Toniutti L., Breitler JC., Etienne H., Campa C., Doulbeau S., Urban L., Lambot C., Pinilla JC., Bertrand B. (2017) Influence of Environmental Conditions and Genetic Background of Arabica Coffee (*C. arabica* L) on Leaf Rust (*Hemileia vastatrix*) Pathogenesis. *Front. Plant Sci*. **8**:2025. doi: 10.3389/fpls.2017.02025.

16. Wrigley G. (1988) *Coffee. Tropical Agriculture Series*. John Wiley and Sons, Inc., New York.

17. International Coffee Organization. http://www.ico.org/trade_statistics.asp accessed at 1st May 2019.

18. Kilambo D., Reuben S., Mamiro D. (2013) Races of *Hemileia vastatrix* and Variation in Pathogenicity of *Colletotrichum kahawae* Isolates to Compact Coffee Genotypes in Tanzania. *J. Plant Stud.* **2**(2): 95-104. doi:10.5539/jps.v2n2p95.

19. Kushalappa AC., Eskes AB. (1989) Advances in Coffee Rust Research. *Annu. Rev. Phytopatol.* **27**:503-531.

20. Bettencourt AJ., Rodrigues Jr., CJ. (1988) *Principles and practice of coffee breeding for resistance to rust and other diseases.* In: Clarke RJ, Macrae R (eds). Elsevier Applied Science, London, **4**:199-234.

21. McCook, S. (2006) Global rust belt: *Hemileia vastatrix* and the ecological integration of world coffee production since 1850. *J. Global Hist.* **1***:* 177–195.

22. Wellman FL. (1970) The rust *Hemileia vastatrix* now firmly established on coffee in Brazil. *Plant Disease Reporter* **54**: 539-541.

23. Avelino J., Cristancho M., Georgiou S., Imbach P., Aguilar L., Bornemann G., Läderach P., Anzueto F., Hruska A.J., Morales C. (2015) The coffee rust crises in Colombia and Central America (2008–2013): Impacts, plausible causes and proposed solutions. *Food Secur.* **7**:303–321. doi: 10.1007/s12571-015-0446-9.

24. Silva M.C., Várzea V., Guerra-Guimarães L., Azinheira H.G., Fernandez D., Petitot AS., Bertrand B., Lashermes P., Nicole M. (2006) Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian Journal of Plant Physiology* **18**(1):119-147. doi:10.1590/s1677-04202006000100010.

25. Hückelhoven R. (2007) Cell wall–associated mechanisms of disease resistance and susceptibility. *Annu. Rev. Phytopathol.* **45:** 101–127. doi:10.1146/annurev.phyto.45.062806.094325.

26. Miedes E., Vanholme R., Boerjan W., Molina A. (2014) The role of the secondary cell wall in plant resistance to pathogens. *Front. Plant Sci.* **5**:358. doi:10.3389/fpls.2014.00358.

27. Jones JDG., Dangl JL. (2006) The plant immune system. *Nature* **444**:323–329. doi:10.1038/nature05286.

28. Bonardi V., Dangl JL. (2012) How complex are intracellular immune receptor signaling complexes?. *Front. Plant Sci.* **3**:1–9. doi:10.3389/fpls.2012.00237.

29. Wit PJGM., Mehrabi R., Van Den Burg HA., Stergiopoulos I. (2009) Fungal effector proteins: Past, present and future. *Mol. Plant Pathol.* **10**: 735–747. doi:10.1111/j.1364-3703.2009.00591.x.

30. Heath MC. (2000) Hypersensitive response-related death. *Plant Mol. Biol* **44**:321–334.

31. Lam E., Kato N., Lawton M. (2001) Programmed cell death, mitochondria and the plant hypersensitive response. *Nature* **411**: 848–853. doi:10.1038/35081184.

32. Silva MC., Guerra-Guimarães L., Loureiro A., Nicole M. (2008) Involvement of peroxidases in the coffee resistance to orange rust (*Hemileia vastatrix*). *Physiol. Mol. Plant Pathol.* **72**(1-3):29-38. doi:10.1016/j.pmpp.2008.04.004.

33. Guerra-Guimarães L., Silva MC., Struck C., Loureiro A., Nicole M., Rodrigues Jr. CJ., Ricardo CPP. (2009) Chitinases of *Coffea arabica* genotypes resistant to orange rust *Hemileia vastatrix*. *Biologia Plantarum* **53**: 702-706.

34. Guerra-Guimarães L., Tenente R., Pinheiro C., Chaves I., Silva MC., Cardoso FMH., Planchon S., Barros DR., Renaut J., Ricardo CP. (2015) Proteomic analysis of apoplastic fluid of *Coffea arabica* leaves highlights novel biomarkers for resistance against *Hemileia vastatrix*. *Front. Plant Sci.* **6:478.** doi: 10.3389/fpls.2015.00478.

35. Guerra-Guimarães L., Vieira A., Chaves I., Queiroz V., Pinheiro C., Renaut J. Silva L., Zambolim L., Ricardo C., Silva MC. (2013) Integrated Cytologic and Proteomic Analysis of *Coffea Arabica – Hemileia vastatrix* Interactions. *ASIC 24th Int. Conf. Coffee Sci.* pp1414-1418.

36. Ilyas M, Rasheed A, Mahmood T. (2016) Functional characterization of germin and germin-like protein genes in various plant species using transgenic approaches. *Biotechnol. Lett.* **38**(9):1405-1421. doi:10.1007/s10529-016-2129-9.

37. Diniz I., Talhinhas P., Azinheira HG., Várzea V., Medeira C., Maia I.; Petitot AS., Nicole M., Fernandez D., Silva MC. (2012) Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and non-host resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. *European Journal of Plant Pathology* **133**:141-157. doi: 10.1007/s10658-011-9925-9.

38. Diniz I., Figueiredo A., Loureiro A., Batista D., Azinheira H., Várzea V., Pereira P., Gichuru G., Moncada P., Guerra-Guimarães L., Oliveira H., Silva MC. (2017) A first insight into the involvement of phytohormones pathways in coffee resistance and susceptibility to *Colletotrichum kahawae*. *PLoS One* **12**:1-20. doi:10.1371/journal.pone.0178159.

39. Diniz I., Azinheira H., Figueiredo A., Oliveira H., Guerra-Guimarães L., Silva MC. (2019) Fungal penetration associated with recognition, signaling and defence-related genes and peroxidase activity during the resistance response of coffee to *Colletotrichum kahawae. Physiol. Mol Plant Pathol.* **105**:119-127. doi:10.1016/j.pmpp.2017.12.005.

40. Ganesh D., Petitot AS., Silva MC., Alary R., Lecouls AC., Fernandez D. (2006) Monitoring of the early molecular resistance responses of coffee (*Coffea arabica* L.) to the rust fungus *(Hemileia vastatrix*) using real-time quantitative RT-PCR. *Plant. Sci.* **170**(6):1045-1051. doi:10.1016/j.plantsci.2005.12.009.

41. Guzzo SD., Harakava R., Tsai SM. (2009) Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. *Journal of Phytopathology* **157**:625-638.

42. Diola V., Brito GG., Caixeta ET. (2013) A new set of differentially expressed signaling genes is early expressed in coffee leaf rust race II incompatible interaction. *Funct. Integr. Genomics* **13:**379. doi:10.1007/s10142-013-0330-7.

43. Barka GD., Caixeta ET., Almeida RF., Alvarenga SM., Zambolim L. (2017). Differential expression of molecular rust resistance components have distinctive profiles in *Coffea arabica-Hemileia vastatrix* interactions. *European Journal of Plant Pathology* **149**(3):543-561. doi:10.1007/s10658-017-1202-0.

44. Silva MC., Nicole M., Guerra-Guimarães L., Rodrigues CJ. (2002) Hypersensitive cell death and post-haustorial defence responses arrest the orange rust (*Hemileia vastatrix*) growth in resistant coffee leaves. *Physiol. Mol. Plant. Pathol.* **60**(4):169-183. doi:10.1006/pmpp.2002.0389.

45. Bustin SA. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **25**(2):169-193. doi:10.1677/jme.0.0250169.

46. Giulietti A., Overbergh L., Valckx D., Decallonne B., Bouillon R., Mathieu C. (2001) An overview of real-time quantitative PCR: Applications to quantify cytokine gene expression. *Methods* **25**(4):386-401. doi:10.1006/meth.2001.1261.

47. Stürzenbaum SR., Kille P. (2001) Control genes in quantitative molecular biological techniques: The variability of invariance. *Comp. Biochem. Physiol. - B. Biochem. Mol. Biol.* **130**(3):281-289. doi:10.1016/S1096-4959(01)00440-7.

48. Kirigia D., Runo S., Alakonya A. (2014) A virus-induced gene silencing (VIGS) system for functional genomics in the parasitic plant *Striga hermonthica*. *Plant Methods* **10**:16.

49. Meng LH., Wang RH., Zhu BZ., Zhu HL., Luo YB., Fu DQ. (2016) Efficient virus-induced gene silencing in *Solanum rostratum*. *PLoS One* **11**(6):1-10. doi:10.1371/journal.pone.0156228.

50. Senthil-Kumar M., Mysore KS. (2014) Tobacco rattle virus-based virus-induced gene silencing in *Nicotiana benthamiana*. *Nat. Protoc.* **9**(7):1549-1562. doi:10.1038/nprot.2014.092.

51. Petrova N., Zenkova M., Chernolovskaya E. (2013) Structure - functions relations in small interfering RNAs. *InTech* **8**:187-227.

52. Ratcliff F., Martin-Hernandez AM., Baulcombe DC. (2001) Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J.* **25**(2):237-245. doi:10.1046/j.0960-7412.2000.00942.x.

53. Liu E., Page JE. (2008) Optimized cDNA libraries for virus-induced gene silencing (VIGS) using tobacco rattle virus. *Plant Methods* **4**(1):1-13. doi:10.1186/1746-4811-4-5.

54. Zhang J., Yu D., Zhang Y., Liu K., Xu K., Zhang F., Wang J., Tan G., Nie X., Ji Q. (2017) Vacuum and co-cultivation agroinfiltration of (germinated) seeds results in Tobacco Rattle Virus (TRV) mediated whole-plant virus-induced gene silencing (VIGS) in wheat and maize. *Front. Plant Sci.* **8**:393. doi:10.3389/fpls.2017.00393.

55. Broderick SR., Jones ML. (2014) An Optimized Protocol to Increase Virus-Induced Gene Silencing Efficiency and Minimize Viral Symptoms in Petunia. *Plant Mol. Biol. Report* **32**:219–233.

56. Shen Z., Sun J., Yao J., Wang S., Ding M., Zhang H., Chen S. (2015) High rates of virus-induced gene silencing by tobacco rattle virus in *Populus*. *Tree Physiology* **35**:1016–1029. https://doi.org/10.1093/treephys/tpv064.

57. Silva MC., Nicole M., Rijo L., Geiger JP., Rodrigues Jr. CJ. (1999) Cytochemical aspects of the

plant-rust fungus interface during the compatible interaction *Coffea arabica* (cv. Caturra)-*Hemileia vastatrix* (race III). *International Journal of Plant Science* **160**: 79-91.

58. Gorgulho R. (2013) *Análise integrada de dados da proteómica e transcriptómica para caracterização da resistência do cafeeiro a Hemileia vastatrix e Colletotrichum kahawae*. pp1-29. Trabalho de Conclusão de Curso (Biologia) - Instituto Superior de Agronomia, Lisboa, 2013.

59. Ramiro DA., Escoute J., Petitot AS., Nicole M., Maluf MP., Fernandez D. (2009) Biphasic haustorial differentiation of coffee rust (*Hemileia vastatrix* race II) associated with defence responses in resistant and susceptible coffee cultivars. *Plant. Pathol.* **58**(5):944-955. doi:10.1111/j.1365-3059.2009.02122.x.

60. Velásquez AC., Chakravarthy S., Martin GB. (2009) Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. *Journal of Visualized Experiments* **28**:1292.doi:10.3791/1292.

61. Liu Q., Yang J., Yan S., Zhang S., Zhao J., Wang W., Yang T., Wang X., Mao X., Dong J., Zhu X., Liu B. (2016) The germin-like protein OsGLP2-1 enhances resistance to fungal blast and bacterial blight in rice. *Plant. Mol. Biol.* **92**: 411–423.

62. Rietz S., Bernsdorff F., Cai D. (2012) Members of the germin-like protein family in *Brassica napus* are candidates for the initiation of an oxidative burst that impedes pathogenesis of sclerotinia sclerotiorium. *Journal of experimental botany* **63**: 5507-5519.

63. Munir F., Hayashi S., Batley J., Naqvi SM., Mahmood T. (2016) Germin-like protein gene-2 promoter from rice is responsive to fungal pathogens in transgenic potato plants. *Funct. Integr. Genomics* **16**(1):19–27.

64. Simões I., Faro C. (2004) Structure and function of plant aspartic proteinases. *Eur. J. Biochem.* **271**:2067–2075.

65. Athauda SB., Matsumoto K., Rajapakshe S., Kuribayashi M., Kubomura-Yoshida N., Iwamatsu A., Shibata C., Inoue H., Takahashi K. (2015) Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. *Biochem. J.* **382**(3):1039-1039. doi:10.1042/bj3821039.

66. Yao X., Xiong W., Ye T., Wu Y. (2012) Overexpression of the aspartic protease *ASPG1* gene confers drought avoidance in *Arabidopsis*. *Journal of experimental botany* **63**(7): 2579–2593. doi:10.1093/jxb/err433

67. Silva J., Scheffler B., Sanabria Y., De Guzman C., Galam D., Farmer A., Woodward J., May G., Oard J. (2012) Identification of candidate genes in rice for resistance to sheath blight disease by whole genome sequencing. *Theor. Appl. Genet.* **124**:63–74.

68. Breen S., Williams SJ., Outram M., Kobe B., Solomon PS. (2017) Emerging Insights into the Functions of Pathogenesis-Related Protein 1. *Trends Plant Sci.* **22**:871–879.

69. Jain S., Srivastava S., Sarin NB., Kav NNV. (2006) Proteomics reveals elevated levels of PR 10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli. *Plant Physiol. Biochem.* **44**:253–259

70. Shiu SH, Bleecker AB. (2003) Plant Receptor-Like Kinase Gene Family: Diversity, Function, and Signaling. *Sci. Signal.* **2001**(113) doi:10.1126/stke.2001.113.re22.

71. Burch-Smith TM. (2006) Efficient Virus-Induced Gene Silencing in Arabidopsis. *Plant. Physiol.* **142**(1):21-27. doi:10.1104/pp.106.084624.

72. Senthil-Kumar M., Mysore KS. (2011) Virus-induced gene silencing can persist for more than 2years and also be transmitted to progeny seedlings in *Nicotiana benthamiana* and tomato. *Plant Biotechnol. J.* **9**(7):797-806. doi:10.1111/j.1467-7652.2011.00589.x.

73. Tian J., Pei H., Zhang S., Chen J., Chen W., Yang R., Meng Y., You J., Gao J., Ma N. (2014) TRV-GFP: A modified Tobacco rattle virus vector for efficient and visualizable analysis of gene function. *J. Exp. Bot.* **65**:311–322.

74. Ramalho TO., Figueira AR., Sotero AJ., Wang R., Geraldino Duarte PS., Farman M., Goodin MM. (2014) Characterization of Coffee ringspot virus-Lavras: A model for an emerging threat to coffee production and quality. *Virology* **464-465**:385-396. doi:10.1016/j.virol.2014.07.031.

75. Boyen C., Oudot MP., Loiseaux-De Goer S. (2001) Origin and evolution of plastids and mitochondria: The phylogenetic diversity of algae. *Cah. Biol. Mar.* **42**(1-2):11-24.

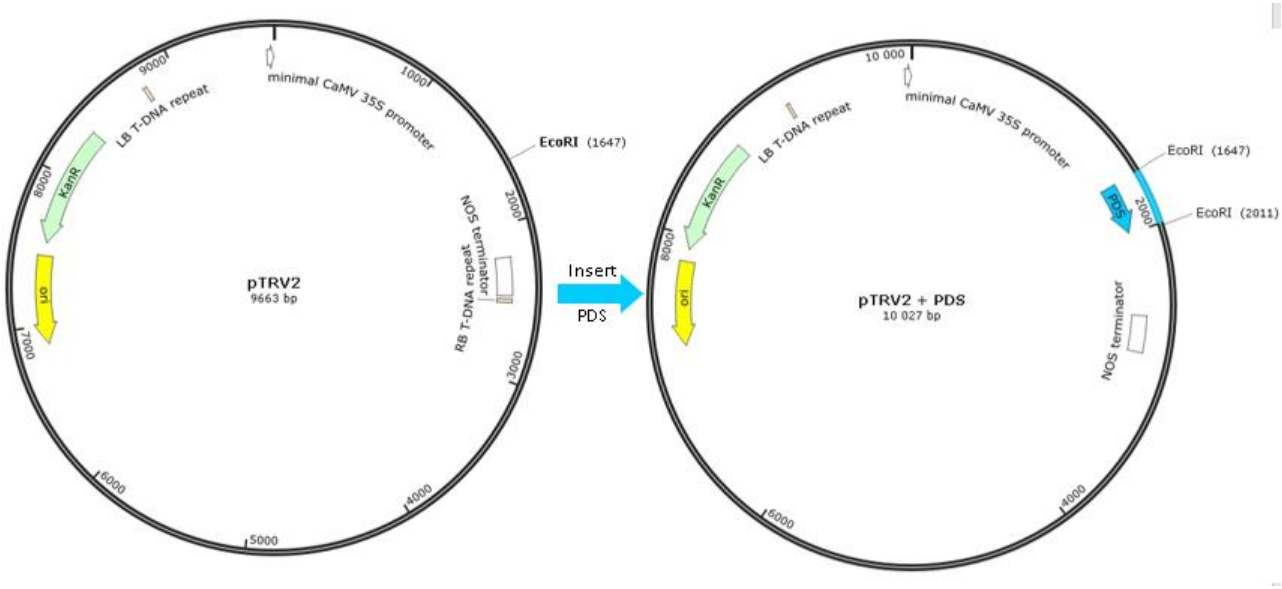
76. Guo X., Ruan S., Hu W., Cai D., Fan L. (2008) Chloroplast DNA insertions into the nuclear genome of rice: The genes, sites and ages of insertion involved. *Funct. Integr. Genomics* **8**(2):101-108. doi:10.1007/s10142-007-0067-2.

77. Goodier JL, Kazazian HH. (2008) Retrotransposons Revisited: The Restraint and Rehabilitation of Parasites. *Cell* **135**(1):23-35. doi:10.1016/j.cell.2008.09.022.

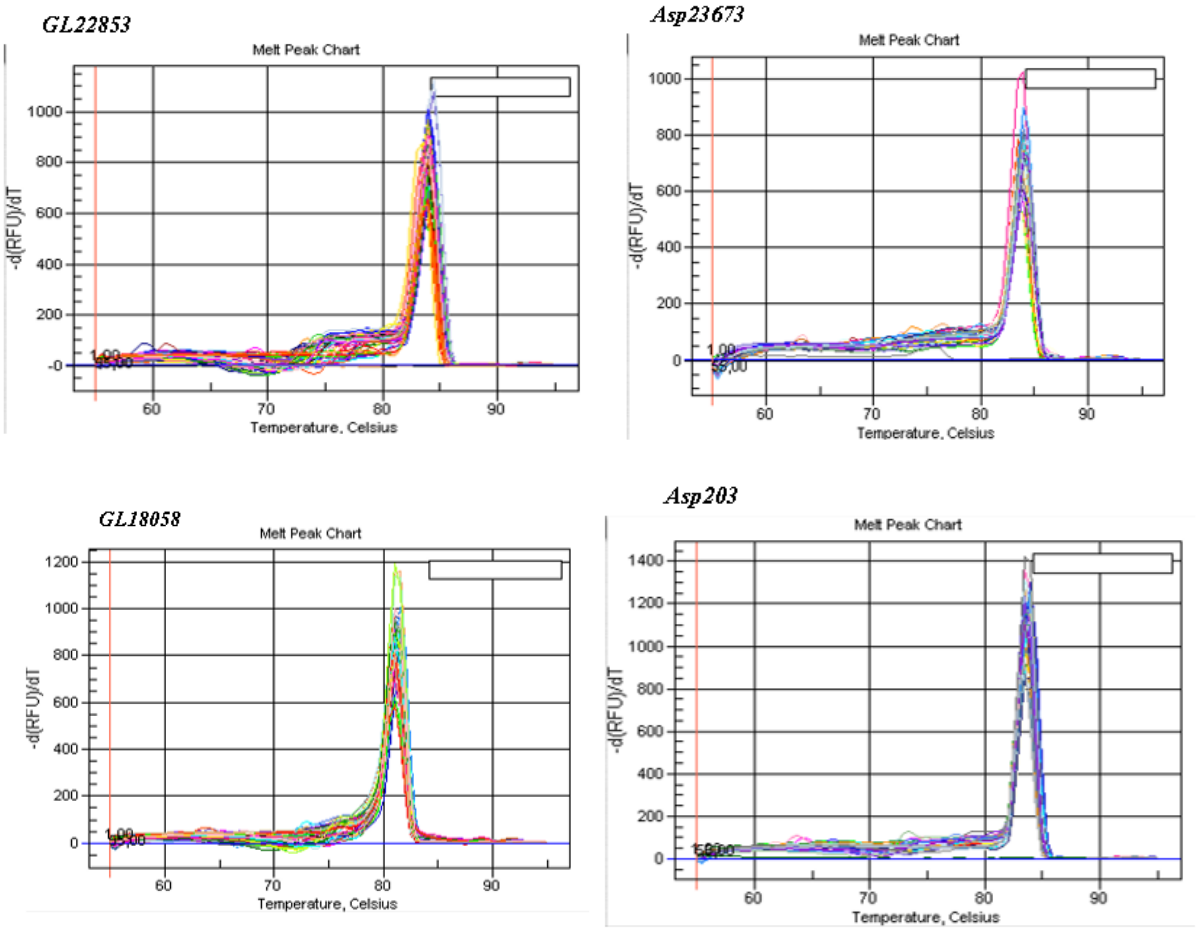
6 SUPPLEMENTARY INFORMATION

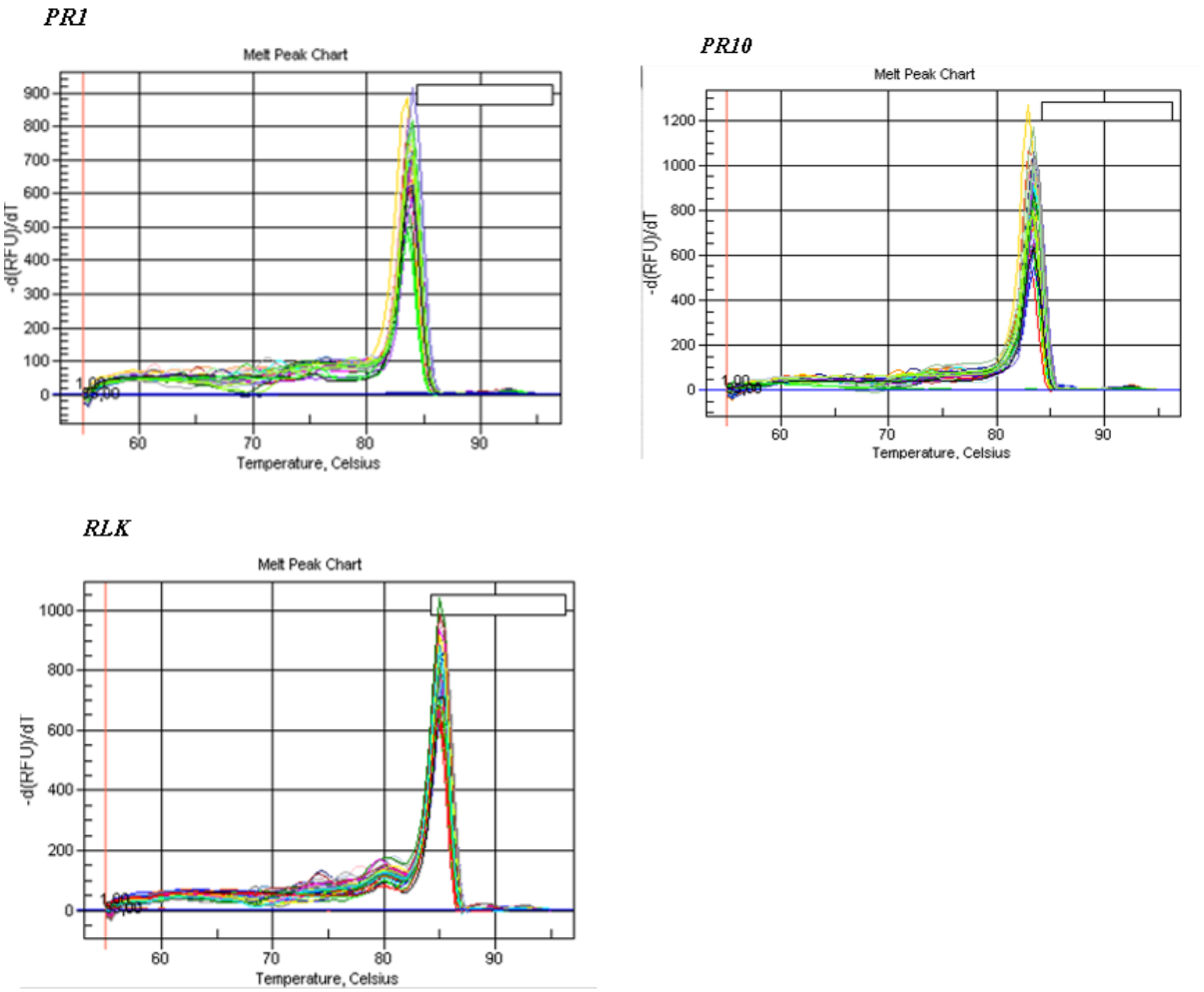
```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        10         20         30         40         50
DQ357179.1      AATATGCAGA ACCTGTTTGG AGAACTAGGA ATTAATGATC GGTTGCAGTG
DQ469932.1      AATATGCAGA ACCTGTTTGG AGAACTAGGG ATTGATGATC GGTTGCAGTG
Consensus       AATATGCAGA ACCTGTTTGG AGAACTAGGR ATTRATGATC GGTTGCAGTG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
DQ357179.1      GAAGGAGCAT TCAATGATAT TTGCAATGCC AAATAAGCCT GGAGAGTTCA
DQ469932.1      GAAGGAACAT TCAATGATAT TTGCGATGCC TAACAAGCCA GGGGAGTTCA
Consensus       GAAGGARCAT TCAATGATAT TTGCRATGCC WAAYAAGCCW GGRGAGTTCA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
DQ357179.1      GTCGATTTGA TTTTCCTGAG GTGCTACCAG CACCATTAAA TGGAATATGG
DQ469932.1      GCCGCTTTGA TTTTCCTGAA GCTCTTCCTG CGCCATTAAA TGGAATTTTG
Consensus       GYCGMTTTGA TTTTCCTGAR GYKCTWCCWG CRCCATTAAA TGGAATWTKG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
DQ357179.1      GCCATCTTGA AGAATAATGA CATGCTTACT TGGCCAGAGA AAGTCAAATT
DQ469932.1      GCCATACTAA AGAACAACGA AATGCTTACG TGGCCCGAGA AAGTCAAATT
Consensus       GCCATMYTRA AGAAYAAYGA MATGCTTACK TGGCCMGAGA AAGTCAAATT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       210        220        230        240        250
DQ357179.1      TGCAATTGGA CTCTTGCCAG CAATTCTGGG TGGACAATCT TATGTTGAGG
DQ469932.1      TGCTATTGGA CTCTTGCCAG CAATGCTTGG AGGGCAATCT TATGTTGAAG
Consensus       TGCWATTGGA CTCTTGCCAG CAATKCTKGG WGGRCAATCT TATGTTGARG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       260        270        280        290        300
DQ357179.1      CACAAGATGG TATAACTGTC AAAGACTGGA TGAGAAAGCA AGGCATACCA
DQ469932.1      CTCAAGACGG TTTAAGTGTT AAGGACTGGA TGAGAAAGCA AGGTGTGCCT
Consensus       CWCAAGAYGG TWTAASTGTY AARGACTGGA TGAGAAAGCA AGGYRTRCCW

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       310        320        330        340        350
DQ357179.1      GATCGGGTGA CTGATGAAGT ATTCTTTGCC ATGTCAAAGG CACTGAACTT
DQ469932.1      GATAGGGTGA CAGATGAGGT GTTCATTGCC ATGTCAAAGG CACTTAACTT
Consensus       GATMGGGTGA CWGATGARGT RTTCWTTGCC ATGTCAAAGG CACTKAACTT
```

Supplementary Fig. 1: Alignment of *PDS* gene DQ357179.1 (*Coffea canephora)* and DQ469932.1 (*Nicotiana benthamiana)* and the respective consensus sequence. Designed primers highlighted in yellow.



Supplementary Fig. 2: TRV2 vector, before and after ligation with *PDS* fragment.





Supplementary Fig. 3: Dissociation curves of studied genes obtained after qPCR amplification.

```
                 ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
Consensus_       AAGCTATTTA GGTGACACTA TAGAATACTC AAGCTATGCA TCCAACGCGT
DQ357179.1       ---------- ---------- ---------- ---------- ----------

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                         60         70         80         90        100
Consensus_       TGGGAGCTCT CCCATATGGT CGACCTGCAG GCGGCCGCGA ATTCACTAGT
DQ357179.1       ---------- ---------- ---------- ---------- ----------

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                        110        120        130        140        150
Consensus_       GATTATGCAG AACCTGTTTG GAGAACTAGG AATTAATGAT CGGTTGCAGT
DQ357179.1       ----ATGCAG AACCTGTTTG GAGAACTAGG AATTAATGAT CGGTTGCAGT

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                        160        170        180        190        200
Consensus_       GGAAGGAGCA TTCAATGATA TTTGCAATGC CAAATAAGCC TGGAGAGTTC
DQ357179.1       GGAAGGAGCA TTCAATGATA TTTGCAATGC CAAATAAGCC TGGAGAGTTC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                       210        220        230        240        250
Consensus_      AGTCGATTTG ATTTTCCTGA GGTGCTACCA GCACCATTAA ATGGAATATG
DQ357179.1      AGTCGATTTG ATTTTCCTGA GGTGCTACCA GCACCATTAA ATGGAATATG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       260        270        280        290        300
Consensus_      GGCCATCTTG AAGAATAATG ACATGCTTAC TTGGCCAGAG AAAGTCAAAT
DQ357179.1      GGCCATCTTG AAGAATAATG ACATGCTTAC TTGGCCAGAG AAAGTCAAAT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       310        320        330        340        350
Consensus_      TTGCAATTGG ACTCTTGCCA GCAATTCTGG GTGGACAATC TTATGTTGAG
DQ357179.1      TTGCAATTGG ACTCTTGCCA GCAATTCTGG GTGGACAATC TTATGTTGAG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       360        370        380        390        400
Consensus_      GCACAAGATG GTATAACTGT CAAAGACTGG ATGAGAAAGC AAGGCATACC
DQ357179.1      GCACAAGATG GTATAACTGT CAAAGACTGG ATGAGAAAGC AAGGCATACC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       410        420        430        440        450
Consensus_      AGATCGGGTG ACTGATGAAG TATTCTTTGC CATGTCAAAG GCACTGAACA
DQ357179.1      AGATCGGGTG ACTGATGAAG TATTCTTTGC CATGTCAAAG GCACTGAAC-

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       460        470        480        490        500
Consensus_      ATCGAATTCC CGCGGCCGCC ATGGCGGCCG GGAGCATGCG ACGTCGGGCC
DQ357179.1      ---------- ---------- ---------- ---------- ----------

                ....|....| ....|....| ....|....| ..
                       510        520        530
Consensus_      CAATTCGCCC TATAGTGAGT CGTATTACAA TT
DQ357179.1      ---------- ---------- ---------- --
```

Supplementary Fig. 4: Sequencing of a fragment of pGEM-Teasy with the insert *CaPDS* amplified with SP6 and T7priemrs. In blue are presented SP6 and T7 primers. In yellow are presented PDS primers.

```
           ....|....| ....|....| ....|....| ....|....| ....|....|
                     10         20         30         40         50
Seq_ca          --CCTTGTGG CTCCCTCGAG GAA-GTTCCC CAGATTGAAA GTACCAGAGA
KY085909.1      GATGTTGTGC TCAGCTTGGA ATACAATCAT GAAGTTGAAA GTACCAGAGA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      60         70         80         90        100
Seq_ca          TTCCTACAGG CATACCATCG CAAAAACTTC CTTGACCGAT TGGATAAATC
KY085909.1      TTCCTAGAGG CATACCATCA GAAAAACTTC CTTGACCGAT TGGATAAATC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
Seq_ca          AAGAAAACAG CGGTAGCAGC TGCAACAGGA GCTGAATATG CAACAGCAAT
KY085909.1      AAGAAAACAG CGGTAGCAGC TGCAACAGGA GCTGAATATG CAACAGCAAT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
Seq_ca          CCAAGGTCGC ATACCCAGAC GGAAACTAAG CTCCCACTCA CGACCCATGT
KY085909.1      CCAAGGTCGC ATACCCAGAC GGAAACTAAG CTCCCACTCA CGACCCATGT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       210        220        230        240        250
Seq_ca          AACAAGCTAC ACCAAGTAAA AAGTGTAGAA CAATTAGTTC ATAAGGACCG
KY085909.1      AACAAGCTAC ACCAAGTAAA AAGTGTAGAA CAATTAGTTC ATAAGGACCG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       260        270        280        290        300
Seq_ca          CCGTTATATA ACCATTCATC AACAGACGCC GCTTCCCATA TTGGGTAAAA
KY085909.1      CCGTTATATA ACCATTCATC AACAGACGCC GCTTCCCATA TTGGGTAAAA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       310        320        330        340        350
Seq_ca          GTGCAAACCT ATAGCTGCAG AAGTAGGAAT AATGGCACCT GAAATAATAT
KY085909.1      GTGCAAACCT ATAGCTGCAG AAGTAGGAAT AATGGCACCT GAAATAATAT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       360        370        380        390        400
Seq_ca          TGTTTCCGTA AAGTAAAGAT CCAGAAACAG GTTCACGAAT ACCATCAATA
KY085909.1      TGTTTCCGTA AAGTAAAGAT CCAGAAACAG GTTCACGAAT ACCATCAATA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       410        420        430        440        450
Seq_ca          TCTACTGGAG GTGCAGCAAT GAAGGCGATA ATAAATACGG AAGTTGCGGT
KY085909.1      TCTACTGGAG GTGCAGCAAT GAAGGCGATA ATAAATACGG AAGTTGCGGT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       460        470        480        490        500
Seq_ca          CAATAAGGTA GGGATCATCA AAACACCAAA CCATCCAATG TAAAGACGGT
KY085909.1      CAATAAGGTA GGGATCATCA AAACACCAAA CCATCCAATG TAAAGACGGT

                ....|....| ....|....| ....|....|
                       510        520        530
Seq_ca          ATTCAGTGCC GCTAGT---A ACCCAGAAA-
KY085909.1      TTTCAGTGCT GGTTATCCAA TTACAGAAGC
```

Supplementary Fig. 5: Alignment of KY085909.1 (CaChl) with sequence obtained from sequencing of the ~700bp fragment.